



La historia secreta
de cómo las redes sociales
empobrecen la mente
y erosionan el mundo

MAX FISHER

LAS REDES DEL CAOS

Índice

Portada

Sinopsis

Portadilla

Dedicatoria

Nota del autor

PRÓLOGO

Consecuencias

1. Atrapados en el casino
2. Todo es como el Gamergate
3. Abriendo el portal
4. La tiranía de los primos
5. Despertar a la máquina
6. El salón de los espejos
7. Las semillas y el viento
8. Las nuevas campanas
9. La madriguera
10. Los nuevos líderes supremos
11. La dictadura de los «me gusta»
12. Infodemia

EPÍLOGO

Denunciar las malas prácticas

Agradecimientos

Notas

Créditos

Visita Planetadelibros.com y descubre
una
nueva forma de disfrutar de la lectura

¡Regístrate y accede a contenidos exclusivos!

Primeros capítulos
Fragmentos de próximas publicaciones
Clubs de lectura con los autores
Concursos, sorteos y promociones
Participa en presentaciones de libros

PlanetadeLibros

Comparte tu opinión en la ficha del libro
y en nuestras redes sociales:



Sinopsis

En 2018, un grupo de investigadores descubrió que el algoritmo de YouTube podía identificar correctamente un vídeo de un niño semidesnudo y determinar que esa característica era el atractivo de la grabación, al tiempo que lo promocionaba entre usuarios que buscaban contenido pedófilo. En Facebook tardaron poco en darse cuenta de que los usuarios pasaban muchas más horas en su página cuando las informaciones que leían sobre vacunas eran alarmistas o conspirativas, con lo que no dudaron en promocionar ese tipo de publicaciones, en especial entre grupos de madres.

Las redes sociales son probablemente el mayor experimento colectivo de la humanidad y su influencia es mucho más profunda de lo que imaginamos. Nuestras mentes, hijos y democracias están a merced de una industria carente de toda ética cuyo único objetivo parece ser apropiarse de nuestro tiempo y atención.

Basándose en años de reportajes en todo el mundo, Max Fisher cuenta la historia de cómo Meta, Twitter, YouTube y otras redes sociales han explotado la fragilidad psicológica de sus usuarios para crear algoritmos que potencien la radicalización, el extremismo y la violencia. Todo ello con un único objetivo: maximizar sus beneficios y la cuenta de resultados de sus empresas.

Las redes del caos

La historia secreta de cómo las redes sociales
empobrecen la mente y erosionan el mundo

Max Fisher

Traducción de Arnau Figueras Deulofeu

ediciones península

*A mi madre y a mi padre, por traerme a este mundo
A Jordan, por estar siempre a mi lado*

Nota del autor

Este libro se basa en entrevistas con cientos de personas que han estudiado o combatido las redes sociales, les han sacado provecho o se han visto afectadas por ellas, así como con empleados y ejecutivos de Silicon Valley. En algunos casos, para facilitar la lectura del texto, el nombre o el título de una fuente pueden aparecer en el apartado «Notas» en lugar de encontrarse en medio del relato. Todas las entrevistas se grabaron salvo una, con un moderador de una empresa subcontratada que pidió que se le mencionara con seudónimo. Revisé nóminas y archivos empresariales que confirman sus declaraciones.

El libro también recurre en gran medida a artículos académicos, actas judiciales y otras muchas fuentes primarias, las cuales aparecen listadas en «Notas» como pruebas que justifican cada cifra y afirmación presentadas en el relato, además de cualquier declaración que no registrara yo mismo. Algunas afirmaciones se basan en investigaciones académicas que aún no se han publicado fuera de este libro. En estos casos, en el apartado «Notas» puede encontrarse un breve resumen de los resultados, la metodología y la autoría de dichos estudios.

PRÓLOGO

Consecuencias

Entrar en la sede central de Facebook puede parecer como entrar en el Vaticano: un centro de poder rodeado de hermetismo y opulencia que abochornaría a cualquier oligarca ruso. La empresa destinó 300 millones de dólares solo a la construcción del número 21,¹ una espaciosa plaza de acero y cristal formada por jardines, patios y restaurantes gratuitos que visité a finales de 2018. Entre reuniones, me llamó la atención un mural de dos pisos de altura que había en un muro trasero, pues me recordó a un famoso artista chino al que poco antes se le había dedicado una exposición en el Museo Guggenheim. Pregunté a la responsable de relaciones públicas que me acompañaba durante la visita si habían pintado el muro a propósito, imitando el estilo de ese artista. Ella sonrió con educación. No se trataba de una imitación. Habían traído al artista en avión expresamente para que pintara un original en los muros de Facebook. Y lo mismo con otros muchos artistas. A todo mi alrededor, programadores con un sueldo magnífico iban ajetreando de aquí para allá por pasillos adornados con murales de un valor incalculable.²

Dentro de mi mochila, entre varias libretas, estaba mi tique de entrada: más de 1.400 páginas de documentos internos, procedentes de zonas de todo el planeta, que ponían al descubierto cómo la mano invisible de Facebook establecía los límites de lo aceptable en política y en los discursos para 2.000 millones de usuarios de todo el mundo. Para el trabajador que me los había filtrado, los archivos eran una prueba de la negligencia y los atajos de la compañía para intentar frenar el creciente revuelo internacional que, según él, los productos de la empresa agravaban o incluso provocaban. Para mí, eran mucho más que eso. Eran una ventana que permitía ver qué opinaban los directivos de Facebook sobre las consecuencias de la aparición y el crecimiento de las redes sociales.

Como otras muchas personas, de entrada yo había supuesto que los peligros de las redes sociales provenían sobre todo del mal uso que

de ellas hacían personajes malvados —propagandistas, agentes extranjeros o vendedores de noticias falsas— y que, en el peor de los casos, las distintas plataformas eran una vía pasiva por la que discurrían problemas que la sociedad ya tenía de antes. Sin embargo, en casi todos los sitios a los que viajé por mi actividad periodística para informar sobre déspotas, guerras y revueltas de países lejanos, se vinculaban una y otra vez hechos extraños y extremos a las redes sociales. Unos disturbios repentinos, una nueva organización radical o la creencia generalizada en alguna estrambótica conspiración: todo ello tenía un vínculo común. Y, a pesar de que en Estados Unidos aún no había estallado la violencia, las similitudes con lo que estaba ocurriendo allí eran innegables. Todas las semanas salía alguna noticia sobre una conspiración de Twitter que llegaba a la política nacional, una subcultura de Reddit que derivaba en el neonazismo o un adicto a YouTube que perpetraba una masacre.

Y la inesperada victoria de Donald Trump en 2016 se había atribuido, en parte, a las redes sociales. Aunque el papel de las plataformas aún se entendía poco, ya era evidente que el ascenso de Trump se debía a unos nuevos y extraños movimientos de base y a unas opiniones muy polarizadas que proliferaban en internet, además de a unos agentes rusos que habían sacado partido de las tendencias de las redes sociales para distorsionar la realidad y reforzar la identidad. Ese patrón mundial parecía indicar algo fundamental de la tecnología, pero nadie era capaz de decirme qué era aquello exactamente, por qué estaba sucediendo o qué significaba.

En el otro extremo del mundo, un joven al que llamaré Jacob —empleado de una de las muchas empresas a las que Silicon Valley subcontrata el trabajo sucio— albergaba más o menos las mismas sospechas que yo. Había hecho sonar todas las alarmas que había podido. Sus jefes le habían escuchado con preocupación, decía, incluso con solidaridad. Habían visto las mismas cosas que él. Había algo en el producto que supervisaban que estaba yendo peligrosamente mal.

Jacob, un chico delgado y amante de los libros, había crecido enamorado de internet y se había pasado años jugueteando con ordenadores. Aquellas tecnologías parecían representar lo mejor de Estados Unidos. Jacob admiraba mucho a magnates de internet como Mark Zuckerberg, director ejecutivo y fundador de Facebook, quien sostenía que conectar el mundo lo convertiría en un lugar mejor.

Cuando Jacob encontró un trabajo en una empresa que revisaba contenido de usuarios de Facebook e Instagram, una de las varias agencias que la compañía subcontrata en todo el mundo, le pareció que empezaba a formar parte de la historia.

Cada día su equipo revisaba miles de publicaciones de todo el mundo y señalaba todas aquellas que incumplieran una norma o se pasaran de la raya. Él tenía la sensación de que era una tarea agotadora pero necesaria. No obstante, a lo largo de algunos meses de los años 2017 y 2018, se habían dado cuenta de que las publicaciones estaban cada vez más cargadas de odio, de conspiraciones y de extremismo. Y notaban que, cuanto más incendiaria era la publicación, más la difundían las plataformas. Les parecía que había un patrón, algo que se repetía en las decenas de sociedades e idiomas que tenían el encargo de supervisar.

Además, creían que su capacidad de frenar el odio y la incitación crecientes se veía limitada por lo mismo que supuestamente tenía que ayudarlos: los montones de reglamentos secretos que dictaban qué podían permitir en las plataformas y que tenían que eliminar. Para los más de 2.000 millones de usuarios de Facebook, estas normas son en su mayor parte invisibles. Están pensadas para que las plataformas sean seguras y civilizadas, y lo regulan todo, desde la frontera entre la libertad de expresión y los discursos de odio hasta los límites de los movimientos políticos permisibles. Pero a medida que los reglamentos se demostraron insuficientes para contener los perjuicios que con frecuencia creaba la propia plataforma, y a medida que la supervisión empresarial de esta parte —muy poco glamurosa— del negocio iba a la deriva, las guías internacionales habían ido ampliándose hasta convertirse en cientos de páginas confusas y a menudo contradictorias. Algunas de las más importantes —sobre cómo identificar el reclutamiento de terroristas o la supervisión de elecciones polémicas— estaban llenas de erratas, errores fácticos y vacíos legales obvios. La chapucería y las lagunas indicaban una peligrosa indiferencia por una labor que Jacob veía como una cuestión de vida o muerte, y eso en un momento en que las plataformas estaban repletas de un extremismo que cada vez más hacía derramar sangre en el mundo real. Tan solo unos meses antes, en Birmania, las Naciones Unidas habían acusado formalmente Facebook de permitir que su tecnología contribuyera a provocar uno de los genocidios más atroces desde la Segunda Guerra

Mundial.

Jacob anotó las conclusiones y las preocupaciones de su equipo para enviarlas a instancias superiores. Pasaron meses. El aumento del extremismo en internet no hizo más que empeorar. Él iba todos los días al trabajo y esperaba delante de su ordenador a que llegase una respuesta de la sede central, en los lejanos Estados Unidos, pero nunca llegó. Jacob tuvo una idea. Consistiría en meterse dentro del sistema de seguridad en el trabajo, extraer a escondidas archivos confidenciales y filtrarlos al extranjero, y luego convencer a los medios de que publicasen sus advertencias en su lugar: todo ello con la esperanza de hacer llegar la información a la pantalla de una persona: Mark Zuckerberg, fundador y director ejecutivo de Facebook. Jacob estaba seguro de que la distancia y la burocracia le impedían llegar a las personas responsables. Con que pudiera hablar con ellos, seguro que querrían arreglar las cosas.

Jacob se puso en contacto conmigo por primera vez a principios de 2018. Una serie de artículos míos en los que había investigado el papel de las redes sociales en el aumento de la violencia masiva en lugares como el pequeño país asiático de Sri Lanka, le parecieron una confirmación de que los problemas que él había observado en la pantalla eran reales. Y que tenían unas consecuencias cada vez mayores y a veces mortíferas. Pero sabía que su palabra por sí sola no sería suficiente. Tendría que sacar a hurtadillas de los ordenadores de su oficina los reglamentos internos de la empresa y los documentos sobre formación. No sería fácil —las máquinas estaban muy bien protegidas y las oficinas estaban monitorizadas de cerca—, pero era posible: un año antes, alguien había logrado hacer llegar algunos de los archivos a *The Guardian* y más adelante también se filtraron documentos a *Vice News*. Jacob creó un programa para sacar los archivos sin ser visto, encriptándolos y eliminando las huellas digitales con las que pudiera encontrarse su rastro o incluso que pudiera conducir hasta el país donde estaban ubicadas sus oficinas. Me envió algunos archivos mediante un servidor seguro. Al cabo de algunas semanas, me subí a un avión para ir a recoger el resto de los archivos y para reunirme con él.

Facebook, al saber lo que había obtenido, me invitó a su elegante sede central y me ofreció la posibilidad de entrevistarme con varios directivos de la empresa. Todos ellos eran profesionales tenaces.

Algunos se habían granjeado una admirable reputación en Washington, en ámbitos como la lucha antiterrorista o la ciberseguridad, antes de sumarse a la fiebre del oro de Silicon Valley. Otros tenían un historial impresionante en el ámbito de los derechos humanos o la política. Nada tenían que ver con los *hackers* de garaje y los soñadores que no habían terminado los estudios que antes gobernaban las plataformas, aunque más adelante se vería a las claras que las ideologías y sesgos de residencia estudiantil de la primera época de Silicon Valley aún se defendían con una convicción casi religiosa en sus campus y seguían integrados en la propia tecnología que difundía esos mismos ideales al resto del mundo.

En mis conversaciones en la sede central de Facebook detecté un extraño patrón. Un ejecutivo me explicaba el desafío que consumía sus días: impedir que los terroristas reclutaran adeptos en la plataforma, ser más listo que los *hackers* de Gobiernos hostiles, determinar qué combinación de palabras constituía una inaceptable incitación a la violencia. Casi cualquier pregunta que formulaba, por sensible que fuese, daba lugar a una respuesta directa y matizada. Cuando un problema estaba por resolver, lo reconocían. Ninguno tuvo que comprobar sus notas para decirme, por ejemplo, cuál era la política de Facebook con respecto a las organizaciones independentistas kurdas o para explicarme sus métodos para elaborar las normas sobre los discursos de odio en tagalo.

Me preguntaba: con unos directivos tan meticulosos y ultracualificados, ¿por qué los problemas para los que articulan unas respuestas tan razonadas parecen ir siempre a peor? Cuando organizaciones pro derechos humanos alertan a Facebook de un peligro inminente salido de su plataforma, ¿por qué tan a menudo la empresa no actúa? ¿Por qué periodistas como yo, con tan poca capacidad de escudriñar las operaciones de la plataforma y una diminuta parte de su personal o presupuesto, siguen publicando informaciones sobre atrocidades y sectas nacidas en Facebook que parecen cogerlos por sorpresa? Pero en algún momento de cada entrevista, cuando les preguntaba por los peligros causados no por personajes malvados que hacen un mal uso de la plataforma, sino por la plataforma en sí, era como si se levantase un muro mental.

«Los abusos que ves no son ninguna novedad —dijo la jefa de políticas internacionales de la compañía cuando le pregunté por las

consecuencias de la plataforma—. Lo que es distinto en este caso es el poder de amplificación de algo como una red social.»³«Como sociedad, aún nos queda mucho para entender todas las consecuencias de las redes sociales»,⁴dijo el jefe de ciberseguridad de la compañía dando a entender que el cambio principal provocado por la tecnología había sido sencillamente reducir la «fricción» en la comunicación, lo cual permitía que los mensajes viajasen más deprisa y más lejos.

Aquella era una imagen extrañamente incompleta de cómo funciona Facebook. En la empresa, muchos parecían casi no ser conscientes de que los algoritmos y el diseño de la plataforma configuran de forma deliberada las experiencias y los incentivos de los usuarios y, por tanto, a los propios usuarios. Estos elementos son el núcleo del producto, la razón por la que cientos de programadores iban de un lado para otro mientras hablábamos. Era como entrar en una fábrica de cigarrillos y que los directivos te dijese que no entendían por qué la gente se quejaba de los efectos sobre la salud de las cajetillas de cartón que ellos vendían.

En un momento dado, mientras hablaba con dos empleados que supervisaban la respuesta ante las crisis, dejé a un lado el modo periodista y los alerté de algo preocupante que había visto. En países de todo el mundo estaba aflorando un rumor espantoso, de forma al parecer espontánea, en Facebook: que unos misteriosos forasteros secuestraban niños locales para convertirlos en esclavos sexuales y extraerles los órganos. Las comunidades expuestas a ese rumor estaban respondiendo de maneras cada vez más peligrosas. Cuando a través de Facebook y WhatsApp el rumor llegó a una zona rural de Indonesia, por ejemplo, en nueve aldeas se habían formado turbas que habían atacado a transeúntes inocentes. Era como si ese rumor fuera algún virus misterioso que convertía comunidades normales en muchedumbres sanguinarias, y aquello parecía salir de la propia plataforma. Los dos empleados de Facebook escucharon y asintieron con la cabeza. Ninguno preguntó nada. Uno comentó de paso que tenía la esperanza de que, algún día, una instancia independiente investigase cosas como aquellas, y luego cambiamos de tema.

En cualquier caso, en Facebook siguieron apareciendo versiones del rumor. Una versión estadounidense, que había surgido primero en el tablón de mensajes 4chan bajo la etiqueta «QAnon», había llegado poco antes a Facebook como una cerilla a un charco de gasolina. Más

adelante, cuando QAnon se convirtió en un movimiento con decenas de miles de seguidores, un informe interno del FBI lo identificó como una amenaza terrorista nacional. Durante todo el proceso, los motores de recomendación de Facebook fomentaron los grupos de QAnon hasta alcanzar enormes cifras de lectores, como si ese no fuera sino otro club más, con lo cual contribuyeron a agrandar la conspiración hasta que tuvo el tamaño de un partido político minoritario, al parecer por ninguna otra razón que los clics continuos que generaba el contenido de QAnon.

No obstante, dentro de los muros pintados de Facebook, el convencimiento de que el producto era algo positivo parecía inquebrantable. El ideal fundamental de Silicon Valley según el cual hacer que la gente pase más y más tiempo conectada a internet enriquecerá su mente y hará del mundo un lugar mejor estaba muy extendido entre los ingenieros que crean y moldean los productos. «A medida que la plataforma llega a más países y más personas, aumentan los riesgos —afirmó una ingeniera veterana sobre el importantísimo canal de noticias de Facebook—. Pero también creo que hay mayores oportunidades para que la gente esté expuesta a nuevas ideas.»⁵Y me aseguró que eliminarían cualquier riesgo derivado de la misión de la plataforma de maximizar la participación de los usuarios.

Más adelante me enteré de que, poco antes de mi visita, varios investigadores de Facebook —designados por la empresa para estudiar los efectos de su tecnología en respuesta a las sospechas crecientes de que el sitio podía estar empeorando las divisiones políticas en Estados Unidos— habían alertado a nivel interno de que la plataforma estaba haciendo justo aquello de lo que los ejecutivos de la compañía, en nuestras conversaciones, habían hecho caso omiso. «Nuestros algoritmos sacan provecho de la atracción del cerebro humano por la división»,⁶ alertaron los investigadores en una presentación de 2018 que luego se filtró al *Wall Street Journal*. De hecho, decía también la presentación, los sistemas de Facebook estaban diseñados de tal forma que ofrecían «un contenido cada vez más divisivo para captar la atención de los usuarios e incrementar el tiempo que estos pasaban en la plataforma». Los ejecutivos guardaron el estudio en un cajón y rechazaron gran parte de las recomendaciones que en él se hacían. Esos consejos habrían consistido en modificar los sistemas de

promoción que elegían lo que ven los usuarios y podrían haber reducido el tiempo que pasaban conectados. La pregunta que yo había traído a los pasillos de Facebook —¿cuáles son las consecuencias de hacer pasar una porción cada vez mayor de la política, la información y las relaciones sociales humanas a través de unas plataformas de internet diseñadas expresamente para manipular la atención?— allí era un absoluto tabú.

Los meses posteriores a mi visita coincidieron con lo que fue el periodo de mayores críticas públicas de la historia de Silicon Valley hasta entonces. Los gigantes de las redes sociales tuvieron que comparecer ante el Congreso estadounidense, se vieron forzados a aceptar leyes extranjeras, multas multimillonarias y amenazas de fragmentación forzosa. Los personajes públicos se referían de forma habitual a las compañías como una de las mayores amenazas de nuestro tiempo. En respuesta a eso, los directivos de las empresas se comprometieron a hacer frente a los daños que emanaban de sus servicios. Inauguraron salas donde se controlaría la integridad de las elecciones y actualizaron sus políticas de revisión de contenidos. Pero su modelo de negocio —mantener a las personas pegadas a sus plataformas cuantas más horas al día mejor— y la tecnología subyacente desplegada para alcanzar este objetivo no cambiaron casi nada. Y mientras los problemas que habían prometido resolver no hacían más que empeorar, ganaban más dinero que nunca.

La nueva década trajo una ola de crisis. La pandemia del coronavirus, conflictos raciales en Estados Unidos, el ascenso meteórico de una nueva extrema derecha violenta y el intento de destrucción de la propia democracia estadounidense. Todas esas crisis pusieron a prueba la influencia de las plataformas en nuestro mundo. O la pusieron al descubierto, ya que sacaron a la luz ramificaciones que llevaban años formándose.

En verano de 2020, una auditoría independiente de Facebook — encargada por la compañía ante la presión de organizaciones de derechos civiles— llegó a la conclusión de que la plataforma era todo lo que sus ejecutivos me habían insistido en que no era. Sus políticas permitían una desinformación descontrolada que podía alterar elecciones. Sus algoritmos y sistemas de recomendación estaban «conduciendo a los usuarios hacia cámaras de resonancia en las que se potenciaba el extremismo»,⁷ lo cual les enseñaba a odiar. Pero quizás

lo peor de todo era que el informe concluía que la empresa no entendía cómo sus productos afectaban a sus miles de millones de usuarios.

Sin embargo, había unas pocas personas que sí lo entendían y, bastante antes de que muchos de nosotros estuviéramos preparados para oírlo, intentaron alertarnos. La mayoría empezaron siendo unos verdaderos creyentes obsesionados con la tecnología, algunos formaban parte de Silicon Valley, motivo por el cual precisamente estaban mejor situados para darse cuenta pronto de que algo estaba yendo mal, para investigarlo y evaluar las consecuencias. Pero las empresas que afirmaban que querían entender esas ideas obstaculizaban sus iniciativas, cuestionaban su reputación y ponían en duda sus conclusiones, hasta que en muchos casos las empresas se vieron forzadas a admitir, aunque fuera solo de forma implícita, que quienes habían hecho sonar las alarmas tenían razón desde el principio. Hicieron su trabajo, al menos de entrada, de forma independiente, empleando métodos muy distintos para abordar la misma pregunta: ¿cuáles son las consecuencias de esta tecnología? Este libro trata sobre la misión de responder a esta pregunta, y está narrado, en parte, a través de las personas que la llevaron a cabo.

Se demostró que la creencia popular inicial de que las redes sociales fomentan el sensacionalismo y la ira, aunque era cierta, se quedaba muy corta. Una cantidad cada vez mayor de pruebas, recogidas por muchos académicos, periodistas, informantes y ciudadanos preocupados, hace pensar que sus efectos son mucho más profundos. Esta tecnología ejerce tal influencia en nuestra psicología y nuestra identidad, y es tan omnipresente en nuestras vidas, que cambia nuestra manera de pensar, comportarnos y relacionarnos unos con otros. El efecto, multiplicado por miles de millones de usuarios, ha sido la transformación de la propia sociedad.

No se puede culpar a Silicon Valley de las flaquezas psicológicas que nos llevan a hacer daño a los demás o a actuar contra nuestros propios intereses. Ni de la profunda polarización cultural, en Estados Unidos y en otras partes, que allanó el camino para convertir esos nuevos espacios en campos para el conflicto partidista, donde se destruye todo sentido compartido de bienestar o realidad. Ni siquiera puede culparse a las compañías más grandes de Silicon Valley del modelo de financiación del sector de la alta tecnología que permitió su

aparición, que consiste en ofrecer inversiones multimillonarias a veinteañeros con dificultades sociales y luego exigirles una rentabilidad inmediata y exponencial, sin que importen mucho los obscenos incentivos que eso genera. Aun así, esas compañías acumularon algunas de las mayores fortunas empresariales de la historia sacando provecho de esas tendencias y debilidades, en un proceso que ha dado lugar a una época totalmente nueva de la experiencia humana. Las consecuencias —aunque, visto desde el presente, parecen casi previsibles al cien por cien, si alguien se hubiese tomado la molestia de prestar atención— quedaban escondidas detrás de una ideología según la cual las personas serían más felices y libres si pasaban más tiempo en internet. Y quedaban ocultas también detrás de una cepa del capitalismo de Silicon Valley que promueve una subcultura de ingenieros contestataria, descarada y casi mesiánica para dirigir las empresas que gobiernan nuestras mentes.

Cuando se presionó a esas compañías para que se comportasen, al menos en parte, como las instituciones de Gobierno *de facto* en las que se habían convertido, terminaron en el centro de crisis políticas y culturales de las que eran culpables parciales. Siendo benévolo, podríamos calificar el arbitraje de una democracia empeñada en su propia destrucción como una labor ingrata: si las compañías no hubieran alcanzado posiciones de tanto poder, no se habrían negado a asumir responsabilidades hasta que se les forzó a hacerlo a punta de pistola regulatoria y, en casi cada etapa del proceso, no habrían puesto en riesgo el bienestar de sus usuarios para seguir generando miles de millones de ingresos mensuales. Puesto que los gigantes de las redes sociales tenían pocos incentivos para afrontar el coste humano de sus imperios —un coste asumido por el resto de las personas, igual que un pueblo situado en el tramo inferior de un río en el que una fábrica desecha residuos tóxicos en su pozo comunitario—, tendrían que hacerlo en su lugar decenas de personas alarmadas externas al sector y desertores de Silicon Valley.

Atrapados en el casino

EL CIELO SE NOS CAE ENCIMA

Renée DiResta tenía a su bebé en el regazo cuando se dio cuenta de que las redes sociales estaban sacando algo peligroso en las personas, algo que ya estaba llegando de forma invisible a su vida y a la de su hijo. Nadie de su círculo inmediato tenía hijos, así que se unió a grupos virtuales para padres primerizos a la búsqueda de consejos sobre métodos para dormir o sobre la dentición. Pero los otros usuarios, aun siendo en su mayoría amables —me dijo— de vez en cuando se enzarzaban en «batallas dialécticas» que se alargaban durante miles de comentarios, todas ellas sobre algo con lo que ella pocas veces se había encontrado en la vida real: las vacunas.

Era el año 2014, y DiResta acababa de llegar a Silicon Valley, donde exploraba *start-ups* para una empresa de inversión. En el fondo seguía siendo analista, algo que hacía desde sus años tanto en Wall Street como, antes de eso, en una agencia de inteligencia que da a entender que era la CIA. Para no perder la agilidad mental, ocupaba su periodo de inactividad con sofisticados proyectos de investigación, al igual que otras personas podrían hacer un crucigrama en la cama.

Como le había picado la curiosidad, empezó a investigar si la furia antivacunas que había visto por internet reflejaba algo más amplio. Se dio cuenta de que, soterrados entre los archivos del departamento de salud pública de California, se recogían los índices de vacunación de casi todos los colegios del estado: también de los jardines de infancia a los que se planteaba llevar a su hijo. Lo que descubrió la asombró. Algunos de los colegios tenían un índice de vacunación de tan solo el 30 %. «¿Qué demonios está pasando?», se preguntó. Se descargó los registros de los diez años anteriores. La tendencia durante ese periodo —un incremento constante de los no vacunados— era clara, me contó. «Mierda —pensó—, esto es terrible.»

Con unos índices tan bajos, los brotes de enfermedades como el

sarampión o la tosferina se convertían en un grave peligro, lo cual ponía en riesgo a los hijos de todos. Llamó a la oficina del senador de su estado para preguntar si se podía hacer algo para mejorar los índices de vacunación. Eso no iba a ocurrir, le dijeron. ¿Realmente eran tan odiadas las vacunas?, preguntó. No, le dijo el funcionario. Sus sondeos indicaban un apoyo de un 85 % a un proyecto de ley que endurecería la obligación de vacunarse en los colegios. Pero los políticos temían al movimiento antivacunas, de una extraordinaria vehemencia —formado por jóvenes padres californianos presas de la paranoia y la indignación— que parecían estar saliendo de Twitter, YouTube y Facebook.

«Aquello fue realmente lo que me hizo caer por esta madriguera —dijo DiResta—. Durante seis meses, no es broma, me dediqué a este tema de las ocho de la mañana a las dos de la tarde.» Con el tiempo, aquella madriguera no la llevó a ninguna mano secreta que estuviera detrás del movimiento antivacunas, sino más bien a las propias redes sociales en las que había surgido. Con la esperanza de organizar parte de aquel 85 % de californianos que apoyaban el proyecto de ley sobre la vacunación, creó un grupo —¿dónde, si no?— en Facebook. Cuando compró anuncios de Facebook para conseguir miembros, se dio cuenta de algo curioso. Al escribir «vacuna» o algo relacionado tangencialmente con este tema en la herramienta de la plataforma para seleccionar el público destinatario de los anuncios, la búsqueda le mostraba grupos y temas que estaban opuestos por completo a las vacunas. Pero eso no era todo: cuando dirigió sus anuncios para que les apareciesen a las madres californianas, las usuarias que los recibieron respondieron con una avalancha de invectivas antivacunas. Era como si en el mundo virtual se hubiesen invertido las opiniones provacunas de su comunidad en la vida real.

Movida por la curiosidad, se unió a varios grupos de Facebook contrarios a las vacunas. Sus usuarios parecían vivir solo para las redes sociales, y hacían circular vídeos de YouTube y coordinaban campañas con etiquetas de Twitter. La mayoría expresaban una auténtica angustia por lo que creían que era una enorme conspiración por poner unas peligrosas inyecciones en los brazos de sus hijos. Pero, si solo representaban el 15 % de los californianos, ¿por qué eran tan dominantes allí? Pronto DiResta se dio cuenta de que Facebook hacía algo extraño: le mandaba una oleada de notificaciones insistiendo en

que siguiera otras páginas antivacunas. «Unirse a un grupo antivacunas —me dijo— era algo transformador.» Casi todas las recomendaciones relacionadas con las vacunas que se le ofrecían eran de contenido antivacunas. «El motor de recomendaciones les daba más y más preponderancia.»

Al poco tiempo, el sistema le recomendó unirse a grupos sobre conspiraciones de otras temáticas. Estelas químicas. Terraplanismo. Y mientras iba hurgando, descubrió otra manera en que el sistema fomentaba la desinformación sobre las vacunas. Al igual que con la herramienta para dirigir los anuncios a un público concreto, escribir «vacunas» en el recuadro de búsqueda de Facebook le mostraba un montón de publicaciones y grupos antivacunas.¹ Aunque las páginas generalistas sobre salud y paternidad a menudo contaban con grupos con muchos más miembros, aquellos resultados aparecían más abajo.

DiResta tenía una sospecha sobre lo que estaba pasando. La habían fascinado los ordenadores desde niña, cuando su padre, un ingeniero biomédico que trabajaba en la investigación sobre el cáncer, la había enseñado a programar a los nueve años. Había tenido una máquina Timex de principios de los ochenta en la que se podía jugar a juegos sencillos. En el instituto, en Nueva York, la ingeniería aunó su amor por la resolución creativa de problemas con los nítidos absolutos de las matemáticas. Hizo unas prácticas en varios laboratorios de resonancias magnéticas, donde ayudó a programar ordenadores para procesar imágenes de escáneres cerebrales.

«Me encantaba la idea de poder crear tu propia solución —dijo—. Me gusta el rigor. Me gusta la lógica.» Y los ordenadores eran una cosa entretenida. Las salas de chat desenfadadas de America Online, el servicio de internet por la línea telefónica, ofrecían emocionantes conexiones aleatorias. Los foros sobre intereses compartidos esotéricos, como la banda favorita de DiResta, Nine Inch Nails, parecían comunidades reales. En la universidad, se graduó en Ingeniería Informática, pero decidió no estudiar un máster y, en cambio, optó por trabajar en el mundo de la inteligencia y las finanzas. No obstante, cuando se asentó el polvo de la crisis financiera, se puso en contacto con amigos que trabajaban en Google. Vente a la Costa Oeste, le dijeron.

Gracias a su trabajo en el sector de la inversión en Silicon Valley centrado en el *hardware*, había aprendido lo suficiente sobre redes

sociales para entender lo que había encontrado en sus búsquedas de Facebook. El motivo por el que el sistema promocionaba tanto a los conspiratorios marginales, terminó dándose cuenta DiResta, era la participación. Las plataformas de redes sociales daban mayor protagonismo al contenido que sus sistemas automatizados habían concluido que maximizaría la actividad *online* de los usuarios, lo cual permitiría a la empresa vender más anuncios. Una madre que acepta que las vacunas son seguras tiene menos motivos para pasar mucho tiempo debatiendo sobre la cuestión en foros virtuales. Los grupos de padres de su mismo parecer a los que se une, aunque tengan muchos miembros, pueden ser relativamente tranquilos. Pero una madre que sospecha que hay una enorme conspiración médica que pone en peligro a sus hijos, según vio DiResta, podía pasarse horas investigando sobre la materia. También era probable que buscara aliados, compartiendo información y coordinando acciones para contraatacar. Para la inteligencia artificial que gobierna la plataforma de una red social, la conclusión es obvia: las madres interesadas en cuestiones sanitarias pasarán mucho más tiempo en internet si se unen a grupos antivacunas. Por tanto, fomentar esos grupos, mediante cualquier método que capte la atención de esas usuarias, hará aumentar su participación en la red social. DiResta sabía que, si no estaba equivocaba, Facebook no solo estaba mimando a los extremistas antivacunas. Los estaba creando.

«Me sentía como Chicken Little, diciéndole a la gente que el cielo se nos caía encima —dijo—. Y ellos me miraban como diciendo: “Pero si no es más que otra publicación en las redes sociales”.» Pero lo que DiResta había percibido era que la plataforma tenía un error estructural. Amigos suyos de Silicon Valley se pusieron en contacto con ella para decirle que estaban viendo alteraciones extrañamente parecidas en todo tipo de comunidades virtuales. Ella detectó que operaba un conjunto de dinámicas comunes, que quizás había incluso un punto de origen común en algún sitio de los intestinos de la web social. Y si aquellos eran los efectos que se daban en algo tan pequeño como las políticas de vacunación de los colegios o los debates sobre videojuegos, ¿qué ocurriría cuando alcanzara el ámbito político o la sociedad más en general?

«Veía el panorama y pensaba: “Esto va a ser un desastre tremendo”», dijo al recordar las circunstancias.

Su viaje terminó llevándola hasta el Estado Islámico y la inteligencia militar rusa. Hasta salas de reuniones del Departamento de Estado e incluso una comparecencia en el Congreso de Estados Unidos. Y hasta una serie de conclusiones asombrosas acerca de la influencia que ejercen las redes sociales sobre todos nosotros. Pero aquello empezó en California, enfrentándose a una corriente radical minoritaria que DiResta aún no sabía que representaba algo mucho más profundo e intratable.

Con una certeza casi absoluta, en Facebook o YouTube nadie tenía la intención de promover la actitud contraria a las vacunas. Los grupos representaban un porcentaje tan minúsculo de sus imperios que los ingresos publicitarios que les reportaban probablemente eran insignificantes. En 2015 Zuckerberg, dando una respuesta implícita al problema, escribió que «la ciencia es de una claridad meridiana: las vacunas funcionan y son importantes para la salud de todos los miembros de nuestra comunidad». Pero la tecnología que estaba formando ese movimiento radical partía de algo que ni siquiera el director ejecutivo de la compañía podía superar: las costumbres culturales y financieras que constituyen el núcleo de su sector.

UNAS GALÁPAGOS ESTADOUNIDENSES

Hace menos de un siglo, el valle de Santa Clara, en el centro de California, era una tranquila extensión de huertas de frutales y fábricas de conservas, salpicado por alguna esporádica grúa Derrick. Esto empezó a cambiar en 1941, cuando la marina japonesa atacó Pearl Harbor, lo cual puso en marcha una serie de hechos que transformaron aquel remanso de paz en una de las mayores concentraciones de riqueza que el mundo haya visto.

La historia de esa transformación —que poco tiene que ver con las leyendas de *hackers* o los relatos de residencias estudiantiles que constituyen la tradición oral de Silicon Valley, en buena medida inventada por sus propios habitantes— infundió en el valle unas características culturales y económicas que se incorporaron a los productos que dominan nuestro mundo cada vez más. Y eso comenzó con una oleada de pioneros que desempeñaron un papel tan crucial como los ingenieros o los directores ejecutivos que vinieron después:

el complejo militar-industrial.

Después de Pearl Harbor, el Pentágono, preparándose para adentrarse en el Pacífico pero temeroso de otro ataque sorpresa, dispersó la producción y la investigación militar por varias zonas de la Costa Oeste que para ellos aún tenían un toque del Lejano Oeste. Una de esas ubicaciones fue Moffett Field, una base aérea en su mayor parte en desuso situada en una bahía protegida, escudada por las montañas de Santa Cruz. Cuando terminó la guerra, la maquinaria bélica se quedó, reconvertida para el enfrentamiento cada vez más tenso con la Unión Soviética. En previsión de una guerra nuclear, el Pentágono animó a los contratistas a trasladar proyectos vitales lejos de los grandes centros de población. El gigante aeroespacial Lockheed obedeció y trasladó su división de misiles y asuntos espaciales al tranquilo valle de Santa Clara, justo detrás del hangar número tres de Moffett Field. Buena parte de la carrera armamentística de la Guerra Fría se llevó a cabo desde sus instalaciones. El cofundador de Apple Steve Wozniak, como muchas personas de su generación, creció viendo cómo su padre iba todas las mañanas a trabajar a Lockheed.

Igual de importante fue un centro de investigación académica nuevo e inusual, situado a unos pocos kilómetros. Frederick Terman, hijo de un profesor de Psicología de la Universidad de Stanford —que entonces no era una gran universidad—, se pasó la Segunda Guerra Mundial en los laboratorios de Harvard supervisando proyectos de investigación compartidos entre el ejército y el mundo académico. Y volvió a casa con una idea: que ese modelo continuase en tiempos de paz, pero que ahora los científicos universitarios colaborasen con empresas privadas. Fundó el Parque de Investigación de Stanford, donde las empresas podían trabajar junto a los investigadores académicos.

Con los contratistas de la Guerra Fría a la vuelta de la esquina, había muchos compradores. Aquella circunstancia atrajo a científicos con talento y a estudiantes de posgrado de la Costa Este, pues se les ofrecía la oportunidad de formar parte de una patente o una *start-up* lucrativas. Los departamentos de investigación de las universidades suelen trabajar duro, al menos en teoría, en nombre del bien común. Stanford desdibujó los límites entre el trabajo académico y el trabajo con ánimo de lucro, una evolución que se convirtió en un elemento central para la visión de Silicon Valley, absorbida y propagada por las

innumerables empresas que pasaban por el Parque de Investigación. El razonamiento era que tener un gran éxito en el negocio tecnológico y hacer progresos en el bienestar humano no solo era compatible, sino que era exactamente lo mismo.

Estas condiciones convirtieron la Santa Clara de los años cincuenta en lo que Margaret O'Mara, destacada historiadora de Silicon Valley, ha llamado unas Galápagos del silicio.² Igual que la peculiar geología y el aislamiento extremo de aquellas islas produjeron especies de pájaros y lagartijas únicas, las peculiares condiciones de este valle californiano dieron lugar a formas de hacer negocios y de ver el mundo que no podrían haber nacido en ninguna otra parte. Y que terminaron dando lugar a Facebook, YouTube y Twitter.

La migración fortuita que fue la semilla de buena parte del ADN tecnológico de Silicon Valley, igual que una iguana a la deriva que terminara en una orilla de las Galápagos, fue un ingeniero cascarrabias llamado William Shockley. En Bell Labs, quizás la más prestigiosa de las empresas de investigación de la Costa Este, Shockley había compartido un Premio Nobel en 1956 por ser un pionero en los nuevos transistores con semiconductores. Aquellos aparatos diminutos, que dirigen o modifican señales eléctricas, son los componentes esenciales de la electrónica moderna. Shockley se convenció de que podía mejorar los métodos de Bell. Cuando la salud de su madre empeoró, volvió a casa, el mismo año que recibió el Nobel, para cuidarla e iniciar su propia empresa de transistores. Su ciudad natal era justamente Palo Alto, a menos de diez kilómetros de Moffett Field. Su diseño del transistor requería sustituir el germanio, el material que se solía usar, por el silicio.

Shockley, que tenía fama de ser un hombre difícil y arrogante, se esforzó para convencer a algunos ingenieros de Bell de que le siguieran. Además, incluso con dinero procedente del Pentágono, pocos científicos con pedigrí querían trasladarse al páramo que era San José. Así que contrató a ingenieros con un historial que limitaba sus oportunidades en Boston: estudiantes que no se habían graduado, inmigrantes o judíos. Algunos, como el propio Shockley, eran brillantes pero era difícil trabajar con ellos. Aquello determinó para siempre la idea de que las empresas emergentes del valle eran un dominio para inadaptados sociales emprendedores que estaban donde

estaban por puros méritos: un legado que conduciría a las futuras generaciones a encumbrar a misántropos que habían dejado los estudios y a excusar culturas empresariales tóxicas, del estilo de Shockley, como parte esencial del modelo. Sin embargo, un año después de la creación de la empresa, perdió todo el talento. Su «propensión a humillar a los empleados»,³ su rechazo instintivo a cualquier idea que no fuera suya y su inclinación por los extremos — tiempo después defendió la eugenesia y afirmó que los negros eran inferiores desde un punto de vista genético— fueron demasiado.

Para los desertores, lo fácil y esperable habría sido traer de vuelta sus innovaciones al este del país, donde aún se encontraba el resto del sector. Pero, en lugar de eso, quizás por ninguna otra razón que el clima californiano, obtuvieron financiación procedente de la Costa Este y se quedaron en el oeste. Como casualmente tenían sus empresas en el valle de Santa Clara, allí se dirigieron también la inversión y el talento futuros del sector de los semiconductores. Ese pequeño sector prosperó gracias a la gran cantidad de ingenieros que ya estaban en la ciudad trabajando para Lockheed, lo cual garantizaba contrataciones de primera categoría para cualquier empresa emergente prometedora. Y el Parque de Investigación de Stanford proporcionaba un fácil acceso a una investigación puntera.

Esa mezcla abundante de talento, dinero y tecnología —los tres ingredientes esenciales— se quedarían en el valle,⁴ y el resto se quedó fuera, mediante una práctica de financiación inhabitual: las inversiones de capital riesgo. Casi no llegó dinero de Wall Street. Los productos eran demasiado misteriosos y el mercado excesivamente opaco para los financieros de otros lugares. Por lo visto, las únicas personas capaces de identificar ideas prometedoras —los propios ingenieros— proporcionaron financiación a las empresas emergentes. Alguien que había hecho algo de dinero con su propio proyecto oía hablar de un nuevo aparatito que se estaba diseñando en la ciudad y ponía capital inicial —capital riesgo— a cambio de un porcentaje de participación en la empresa.

Aquella situación iba más allá del dinero. Un inversor eficaz de capital riesgo, para proteger su inversión, a menudo ocupaba una silla en el consejo de administración de la empresa, ayudaba a seleccionar al equipo ejecutivo o incluso hacía de mentor del fundador en persona. Y los inversores de capital riesgo tendían a financiar a

personas en las que confiaban, es decir, individuos a los que conocían personalmente o que se les parecían y hablaban como ellos.⁵ Eso significó que cada generación de ingenieros exitosos cosificó sus puntos fuertes, sus sesgos y puntos ciegos en la siguiente, como una especie aislada cuyas características se vuelven más pronunciadas con cada generación.

A medida que los semiconductores evolucionaron para formar el circuito impreso, luego el ordenador, luego internet y luego las redes sociales, cada tecnología produjo varias figuras estelares de éxito, quienes a su vez financiaron y guiaron a la siguiente cuadrilla de personajes exitosos. Durante todo el proceso, esa comunidad siguió siendo unas Galápagos comerciales-culturales, libres para desarrollar sus propias prácticas hiperespecíficas sobre cómo debe funcionar un negocio, qué constituye el éxito y qué responsabilidades tiene una empresa ante sus consumidores y el mundo en general.

Las consecuencias de ese modelo, con todas sus particularidades, no serían evidentes hasta que los sucesores de Shockley asumieron, en la forma de los gigantes de las redes sociales, un control indirecto sobre todos nosotros. Pero ya se vislumbraban las primeras señales a mediados de la década del año 2000, cuando Silicon Valley empezó a jugar con un elemento de *hardware* mucho más complejo que cualquier semiconductor u ordenador: la mente humana.

CONTRA EL CANAL DE NOTICIAS

Si tuviésemos que ubicar el inicio de la era de las redes sociales, podríamos escoger septiembre de 2006, cuando los operadores de un sitio web creado en una residencia estudiantil —Facebook.com— hicieron un descubrimiento accidental mientras trataban de resolver un problema empresarial. Desde que habían puesto en funcionamiento el sitio dos años y medio antes, habían hecho una entrada moderadamente exitosa en el sector moderadamente exitoso de las redes sociales, en el que los usuarios tenían páginas con perfiles personalizados y no hacían mucho más. En ese momento, Facebook tenía ocho millones de usuarios,⁶ una cantidad impresionante para una pandilla de chavales que apenas tenían la edad reglamentaria para beber alcohol, pero aún no podían garantizar su supervivencia. Incluso

Friendster, que ya se veía entonces como un fracaso estrepitoso, tenía unos diez millones. Lo mismo que LiveJournal. Orkut tenía quince millones. Y Myspace se acercaba a los cien millones.

Las dos ventajas competitivas de Facebook se percibían como inconvenientes. Su diseño austero lo hacía visualmente atractivo, pero menos lucrativo que LiveJournal, repleto de anuncios, o Myspace. Y el hecho de ser exclusivo para los campus universitarios le había proporcionado una gran cuota en un mercado que era limitado y con pocos recursos. La empresa había intentado expandirse a los lugares de trabajo, pero pocos empleados se habían unido. ¿Qué adulto con un mínimo de amor propio colgaría su vida profesional en una web para universitarios?

El crecimiento de los usuarios se había estancado cuando, ese verano, apareció un bote salvavidas: Yahoo ofreció mil millones de dólares por la compra de Facebook. El gigante de internet generaba como mínimo esa misma cantidad de ingresos cada trimestre. Pero su negocio como portal web iba quedando obsoleto y la compañía estaba buscando nuevos mercados para crecer. Las redes sociales parecían algo prometedor. Pero para sorpresa de gran parte del sector, tras meses de negociaciones, Zuckerberg rechazó la oferta. No quería bajarse de la montaña rusa de las *start-ups* y, a los veintidós años, convertirse en un mero trabajador de una Yahoo anquilosada y pasada de moda.⁷No obstante, negar a los empleados que se pasaban noches enteras trabajando la oportunidad de jubilarse en la veintena puso a Zuckerberg bajo una presión tremenda, no solo para transformar Facebook, sino para alcanzar un éxito tan apabullante que los mil millones de Yahoo parecieran calderilla.

La parte dos de su plan de dos partes era terminar abriendo Facebook a cualquier persona. Pero la fallida expansión a los lugares de trabajo generaba incertidumbre sobre su posible éxito e incluso podía ser contraproducente si ahuyentaba a los universitarios, motivo por el cual se confiaba tanto en la parte uno. Zuckerberg reformaría la página de inicio de Facebook para que cada usuario viera un canal de noticias personalizado que incluiría lo que sus amigos estuvieran haciendo en el sitio. Hasta entonces, uno tenía que visitar cada perfil o grupo manualmente para ver la actividad de los demás. Ahora, si un amigo cambiaba el estado de su relación, otro hacía una publicación sobre lo mala que era la pizza en la cafetería u otro se apuntaba a un

evento, todo aquello aparecería en la página de inicio.

Esa retahíla de novedades tenía un nombre: *news feed* o canal de noticias. Fue presentado como una fiesta permanente en la que asistiesen todas las personas que conocías. Pero a algunos usuarios les daba la sensación de que los metían a la fuerza en un panóptico, donde todo el mundo tenía una visibilidad total y absoluta de la vida digital del resto de las personas. Aparecieron de la nada montones de grupos de Facebook con nombres como «Estudiantes contra el canal de noticias de Facebook». En los grupos no pasaba nada tangible. Unirse señalaba que uno estaba de acuerdo; eso era todo. Pero, gracias al rediseño del sitio, cada vez que alguien se unía, todos los amigos de esa persona recibían una notificación en su canal de noticias informándoles de ello. Con un clic del ratón, podían unirse ellos también, de lo cual, a su vez, sus amigos serían informados. En pocas horas, los grupos estaban por todas partes. Uno atrajo a 100.000 miembros en su primer día y a finales de semana ya tenía casi un millón.

En realidad, solo se unió una parte pequeña de los usuarios. Pero la proliferación de las actualizaciones hacía que pareciesen una mayoría aplastante. Y el canal de noticias convertía cada perezoso clic del botón «Unirse» en un grito apasionado: «Contra el canal de noticias» u «ODIO FACEBOOK». La apariencia de una indignación generalizada era, pues, una ilusión. Pero los instintos humanos para adaptarse a lo que hacen los demás están muy arraigados. Los psicólogos han descubierto que, cuando las personas piensan que hay consenso en algo, tienden no solo a aceptarlo, sino a interiorizar ese sentimiento como propio.⁸

Pronto la indignación se volvió acción. Decenas de miles de personas enviaron emails al servicio de atención al cliente de Facebook. A la mañana siguiente, camiones de televisión por satélite asediaron las oficinas de Facebook en Palo Alto, y lo mismo hicieron tal cantidad de manifestantes que la policía pidió a la empresa que se plantease dar marcha atrás en lo que fuera que hubiese causado tal polémica. Dentro de Facebook, algunos estuvieron de acuerdo. La crisis se apaciguó a nivel externo con unas malhumoradas disculpas públicas de Zuckerberg —«Tranquilizaos. Respirad. Os estamos escuchando»—⁹e, internamente, llegando a una conclusión irónica: la ira la estaba creando el mismo producto de Facebook contra el cual

los usuarios estaban protestando.

La amplificación digital había hecho creer erróneamente a los usuarios de Facebook, e incluso a sus directivos, que las voces más ruidosas de la plataforma representaban a todo el mundo, lo que convirtió una pequeña llamarada de indignación en un incendio. Pero lo significativo es que también había hecho otra cosa: había hecho aumentar la participación en la plataforma. Y mucho. En un sector en el que la participación de los usuarios es el principal indicador de éxito, y en una empresa deseosa de demostrar que detrás del rechazo de la oferta multimillonaria de Yahoo había algo más que soberbia, las distorsiones del canal de noticias no solo eran toleradas sino aceptadas gustosamente. Pronto Facebook permitió que cualquier persona crease un perfil en la plataforma. Los índices de crecimiento de los usuarios, que apenas se habían alterado durante la anterior ronda de expansión, se dispararon entre un 600 y un 700 %. El tiempo medio que cada persona pasaba conectada también se incrementó con rapidez. Tan solo trece meses después, en otoño de 2007, la compañía estaba valorada en 15.000 millones de dólares.¹⁰

Creo que ese fue el momento «monolito» de Silicon Valley, parecido a la escena del principio de *2001: Una odisea del espacio*, de Stanley Kubrick, cuando aparece una columna negra delante de un clan de chimpancés, quienes de repente aprenden a manejar herramientas. Aquel avance permitió a Facebook situarse un paso por delante de competidores que antes lo superaban de mucho. Otros desaparecieron cuando, en su lugar, emergió una nueva generación.

En 2006, cuando se introdujo el canal de noticias, un 11 % de los estadounidenses estaban en las redes sociales.¹¹ Entre un 2 y un 4 % eran usuarios de Facebook. Menos de diez años después, en 2014, casi dos tercios de los estadounidenses usaban alguna red social, entre las cuales Facebook, YouTube y Twitter eran casi universales. Ese año, en medio de la segunda legislatura de Obama, se cruzó un umbral significativo en la experiencia humana. Por primera vez, los doscientos millones de estadounidenses con una cuenta activa en Facebook pasaron en promedio más tiempo en la plataforma (cuarenta minutos al día) que socializando en persona (treinta y ocho minutos). Solo dos años después, en verano de 2016, casi un 70 % de los estadounidenses usaban alguna plataforma propiedad de Facebook, a una media de cincuenta minutos al día.¹²

Esos sistemas enganchaban a tantos usuarios con tanta eficacia que, para entonces, el valor de mercado de Facebook —un servicio virtual gratuito con casi ningún producto físico ni servicio al consumidor— superó el de Wells Fargo, uno de los bancos más grandes del mundo. Ese mismo año también superó a General Electric y a JPMorgan Chase; y a finales de 2017 a ExxonMobil. Desde entonces, dos de las mayores compañías del mundo han sido Facebook y Google, otro servicio virtual en su mayoría gratuito que obtiene gran parte de sus ingresos gracias a la publicidad, sobre todo en YouTube, su empresa subsidiaria.

Mucho tiempo después de que fuera evidente el potencial dañino de su tecnología, las compañías afirmarían que ellas solo satisfacen — y nunca moldean o manipulan— los deseos de los usuarios. Pero la manipulación había estado incorporada en los productos desde los inicios.

EL EFECTO CASINO

«Cuando Facebook daba sus primeros pasos, había gente que se me acercaba y me decía: “Yo no estoy en las redes sociales” —recordaba años más tarde Sean Parker, que se había convertido en el primer presidente de Facebook a los veinticuatro años—. Y yo les decía: “Pues, bueno, ya estarás”. Y luego ellos respondían: “No, no, no. Yo valoro mis interacciones en la vida real. Yo valoro los instantes. La presencialidad. La intimidad”. Y yo decía: “Al final vas a caer”.»¹³

Parker estaba orgulloso de ser un *hacker*, al igual que buena parte de la generación de Silicon Valley surgida en los años noventa, cuando el término aún denotaba una suerte de actitud contracultural guay. De hecho, la mayoría creaban *software* para empresas. Pero Parker había sido cofundador de Napster, un programa de compartición de archivos cuyos usuarios distribuían tanta música pirateada que, cuando los pleitos pusieron fin a su actividad dos años después de su introducción, había herido de muerte el negocio de la música. Parker sostenía que había forzado al sector a evolucionar sacando provecho de su apatía ante el mundo digital. Muchos de los artistas y ejecutivos, no obstante, lo veían como un parásito.

La estrategia de Facebook, como la describió él, no estaba muy

alejada de la de Napster. Pero, en lugar de sacar provecho de los puntos débiles del sector musical, haría lo mismo con la mente humana. «El razonamiento que condujo a la creación de esas aplicaciones —dijo Parker en rueda de prensa— era “¿Cómo podemos consumir la mayor cantidad posible de tu tiempo y de tu atención consciente?”» Para conseguirlo, dijo, «de alguna forma tenemos que darte un pequeño chute de dopamina de vez en cuando porque alguien haya puesto un “me gusta” o haya comentado una foto, una publicación o lo que sea. Y eso va a conseguir que aportes más contenido, y eso te va a dar más “me gusta” y comentarios». A eso lo llamó el «bucle de retroalimentación por validación social» y dijo que era «justo lo que se le ocurriría a un *hacker* como yo, porque estás sacando provecho de una vulnerabilidad de la psicología humana». Él y Zuckerberg «entendieron eso» desde el principio, dijo, y «aun así, lo hicimos».

En Silicon Valley, el aprovechamiento de esa vulnerabilidad, lejos de ser un secreto oculto, se debatía abiertamente como una herramienta fascinante para el crecimiento empresarial. El término es «persuasión»: entrenar a los consumidores para alterar su comportamiento de maneras que sean beneficiosas para el balance de cuentas. La Universidad de Stanford tenía operativo desde 1997 su Persuasive Tech Lab, un laboratorio sobre persuasión. En 2007, los ingresos publicitarios derivados de los proyectos de los estudiantes de un solo semestre fueron de un millón de dólares.¹⁴

«¿Cómo unas empresas que producen poco más que bits de código mostrados en una pantalla logran controlar, al parecer, las mentes de los usuarios?», se preguntaba Nir Eyal, destacado asesor de producto de Silicon Valley, en su libro, publicado en 2014, *Enganchado (hooked): cómo construir productos y servicios exitosos que formen hábitos*. «Nuestras acciones han sido manipuladas», explicaba. Servicios como Twitter o YouTube «alteran de forma habitual nuestra conducta cotidiana, justo como pretendían quienes los diseñaron».

Uno de los modelos favoritos de Eyal es la máquina tragaperras. Está diseñada para dar a todas tus acciones una respuesta visual, acústica y táctil. Un chin cuando insertas una moneda. Un crac cuando tiras de la palanca. Un destello de luz de colores cuando la sueltas. Eso se conoce como condicionamiento de Pávlov, en honor al fisiólogo ruso Iván Pávlov, que hacía sonar una campanilla cada vez que daba

de comer a su perro hasta que al final el simple chin-chin de la campanilla despertaba el estómago y las glándulas salivales del animal, como si ya no supiera diferenciar el tañido de una campana de la sensación física de hambre. Las máquinas tragaperras funcionan de esa misma forma: entrenan la mente para que mezcle la emoción de ganar con sus soniditos mecánicos. El acto de tirar de la palanca, que antes no tenía ninguna importancia, se convierte en algo placentero por sí solo.

Eso se debe a una sustancia química neurológica llamada dopamina, la misma que Parker mencionó en aquella rueda de prensa. El cerebro segrega pequeñas cantidades de esta sustancia cuando satisfacemos alguna necesidad básica, ya sea biológica (el hambre o el sexo) o social (el afecto o la validación). La dopamina crea una asociación positiva con cualquier comportamiento que induzca su segregación, lo cual nos enseña a repetir esas conductas. Pero cuando abusamos de ese sistema de recompensa mediante la dopamina, eso puede llevarnos a repetir conductas autodestructivas. A apostar una y otra vez, a emborracharnos o a pasar horas en aplicaciones aunque nos hagan infelices.

La dopamina es el cómplice de las redes sociales dentro de nuestro cerebro.¹⁵ Es el motivo por el cual tu móvil parece una máquina tragaperras, con sus notificaciones de colorines, sus zumbidos y sus suaves vibraciones. Esos estímulos no tienen ningún significado neurológico por sí solos. Pero el móvil los vincula a actividades, como mandar un mensaje a un amigo o mirar fotos, que son gratificantes por naturaleza.

Las aplicaciones sociales se apropian de una compulsión —la necesidad de conectar con otras personas— que puede ser más poderosa que el hambre o la avaricia. Eyal describe a una mujer hipotética, Barbra, que se conecta a Facebook y ve una foto que ha colgado un familiar suyo. Cuando va clicando para ver más fotos o comentarios de respuesta, su cerebro vincula el sentirse conectado con personas a las que quiere con los pitidos y destellos de la interfaz de Facebook. «Con el tiempo —escribe Eyal—, Barbra empieza a asociar Facebook con su necesidad de conexión social.» Aprende a satisfacer esa necesidad con una conducta —usar Facebook— que, en realidad, raras veces se la satisfará.

Poco después del descubrimiento del canal de noticias de

Facebook, las grandes plataformas de redes sociales convergieron en lo que Eyal ha llamado uno de los secretos más poderosos del casino: el refuerzo variable intermitente. El concepto, aunque suene esotérico, es de una tremenda simplicidad. El psicólogo B. F. Skinner descubrió que si asignaba a un sujeto humano una tarea repetitiva —por ejemplo, resolver un rompecabezas sencillo— y le daba una recompensa cada vez que lo terminaba, el individuo solía cumplir las órdenes, pero se detenía justo después de que el psicólogo dejaba de darle recompensas. Pero, si solo le entregaba una recompensa algunas veces y con un tamaño aleatorio, entonces el sujeto completaba la tarea de una forma mucho más sistemática, incluso con empecinamiento. Y seguía realizando la tarea durante mucho más rato una vez que ya se habían terminado las recompensas, como si persiguiera de un modo compulsivo la posibilidad de obtener una recompensa.¹⁶

Las máquinas tragaperras aprovechan esta debilidad psicológica para conseguir un efecto increíble. La impredecibilidad de sacar un premio hace que sea difícil parar. Pues las redes sociales hacen lo mismo. Hacer una publicación en Twitter podría generar un gran premio social, en forma de «me gusta», retuits o respuestas. O podría no generar ninguna recompensa. No saber nunca el resultado hace que sea más difícil dejar de tirar de la palanca. El refuerzo variable intermitente es una característica definitoria no solo del juego, las apuestas y las adicciones, sino también —y eso es algo revelador— de las relaciones abusivas. Los abusadores pasan de forma impredecible de la amabilidad a la crueldad y castigan a sus parejas por comportamientos que antes habían recompensado con afecto. Eso puede conducir a algo llamado vínculo traumático. El miembro de la pareja victimizado busca compulsivamente una respuesta positiva, al igual que un jugador inserta monedas en la máquina o un adicto a Facebook es incapaz de salir de la plataforma; aunque a muchos, eso únicamente les hace sentirse más solos.

Además, mientras que hacer publicaciones en las redes sociales puede parecer una interacción auténtica entre tú y un público, hay una diferencia fundamental invisible. En el mundo virtual, la plataforma actúa como un intermediario inadvertido. Decide cuáles de tus comentarios distribuye a qué usuarios y en qué contexto. Tu próxima publicación podría aparecerles a personas a quienes les

encantará y que te aplaudirán o a personas que la odiarán y te acosarán con preguntas, o ni lo uno ni lo otro. Nunca lo sabrás porque sus decisiones son invisibles. Lo único que sabes es que oyes vítores, abucheos o silencios.

A diferencia de las máquinas tragaperras, las cuales no solemos tener a mano en nuestra vida cotidiana, las aplicaciones de redes sociales son de los productos más accesibles del mundo. Es un casino que cabe en tu bolsillo, por lo cual nos entrenamos poco a poco a responder a cualquier disminución de nuestra felicidad tirando de la máquina tragaperras más omnipresente de la historia. El estadounidense medio consulta el móvil ciento cincuenta veces al día,¹⁷ a menudo para abrir alguna red social.¹⁸ No lo hacemos porque comprobar compulsivamente las redes sociales nos haga felices. En 2018, un equipo de economistas ofreció a varios usuarios distintas cantidades de dinero para que desactivasen su cuenta durante cuatro semanas.¹⁹ Querían encontrar el umbral a partir del cual por lo menos la mitad dijera que sí. La cifra resultó ser elevada: 180 dólares. Pero las personas que desactivaron su cuenta experimentaron más felicidad, menos ansiedad y mayor satisfacción vital. Una vez terminado el experimento, usaron la aplicación menos que antes.

¿Por qué a esos individuos les había costado tanto renunciar a un producto que los hacía infelices? Su comportamiento, escribieron los economistas, era «coherente con los modelos estándar de creación de hábitos» —es decir, con las adicciones— que conducen a «elecciones de consumo no óptimas». Una manera clínica de decir que a los sujetos se les había enseñado a actuar en contra de sus propios intereses.

EL SOCIÓMETRO

Un año después de introducir el canal de noticias, un grupo de programadores de Facebook creó un prototipo de algo que llamaron el «botón de genial»: expresaría aprobación respecto a la publicación de otro usuario con un solo clic. Zuckerberg rechazó la idea varias veces, pues creía que alejaría a los usuarios de otros comportamientos más interesantes, como publicar comentarios. Empezó a «considerarse un proyecto maldito porque en muchas revisiones de Zuck no había

pasado el corte»,²⁰ escribió Andrew Bosworth, uno de los desarrolladores del canal de noticias, que más adelante llegó a ser vicepresidente de Facebook. Tras un año y medio en el limbo, un nuevo equipo se hizo cargo de lo que ahora era el botón de «like» o «me gusta». En algunas pruebas que hicieron con usuarios, escribió Bosworth en una publicación recordando el episodio, vieron que el botón incrementaba el número de comentarios. Al verlo, Zuckerberg cedió.

A principios de 2009, una jefa de producto llamada Leah Pearlman, que había trabajado en aquella función desde poco después de llegar a Facebook a los veintitrés años, hizo una publicación anunciando que era «una manera sencilla de decir a los amigos que te gusta lo que han compartido en Facebook con un solo clic». El tráfico aumentó de inmediato, muy por encima de las expectativas internas. Pero el comportamiento de los usuarios también cambió. Para todos los Nir Eyal y los Sean Parker que hablan sobre usuarios adictos, aquello fue, al igual que el canal de noticias y otras muchas evoluciones que estaban por llegar, otro caso en el que las empresas de redes sociales dieron con un truco psicológico aún más potente que no entendían.

El atractivo de ese botoncito y buena parte del poder de las redes sociales reside en el hecho de sacar partido a algo llamado el sociómetro. El concepto apareció a raíz de una pregunta formulada por el psicólogo Mark Leary: ¿cuál es el propósito de la autoestima? La angustia que sentimos por culpa de la baja autoestima es totalmente autogenerada. No habríamos desarrollado una vulnerabilidad tan inusual y dolorosa, sostenía Leary, a menos que nos proporcionase algún beneficio que superara sus terribles costes psíquicos. Su teoría, ahora ampliamente aceptada, es que la autoestima es, de hecho, «un termómetro psicológico del grado en que las personas perciben que son valoradas en términos relacionales y aceptadas socialmente por los demás».²¹

Los seres humanos somos uno de los animales sociales más complejos de nuestro planeta. Evolucionamos para vivir en colectivos sin líderes mucho mayores que los otros primates: de hasta unos 150 miembros.²² Como individuos, nuestra capacidad de prosperar dependía de lo bien que gestionábamos esas 149 relaciones: por no decir las relaciones de todos los seres de nuestro alrededor entre ellos.

Si el grupo nos valoraba, podíamos confiar en que tendríamos apoyo, recursos y probablemente un individuo con quien aparearnos. De lo contrario, podíamos no conseguir ninguna de esas cosas. Era cuestión de supervivencia, desde un punto de vista físico y genético.

A lo largo de millones de años, esas presiones seleccionaron individuos que son sensibles y hábiles en la maximización de su estatus. Es lo que el antropólogo Brian Hare denominó «la supervivencia de los más simpáticos».²³ El resultado fue el desarrollo de un sociómetro: una tendencia a monitorizar de forma inconsciente cómo los otros individuos de nuestra comunidad parecen percibirnos. Procesamos esta información a modo de autoestima y de emociones relacionadas, como el orgullo, la vergüenza o la inseguridad.²⁴ Estas emociones nos empujan a realizar más acciones de las que hacen que nuestra comunidad nos valore y menos de las que no. Y, de un modo significativo, están pensadas para hacer que la motivación parezca que nos sale de dentro. Si nos diéramos cuenta, a un nivel consciente, de que estamos respondiendo a la presión social, nuestra actitud podría parecernos reticente o cínica, lo que la haría menos convincente.

La función de los «me gusta» de Facebook, de la que ahora existe una versión similar en todas las plataformas, es el equivalente de una batería de coche enganchada a ese sociómetro. A quien controla las descargas eléctricas le da un poder enorme sobre nuestro comportamiento. No es solo que los «me gusta» proporcionen la validación social que dedicamos tantas energías a conseguir; es que la ofrecen con una inmediatez y una magnitud desconocidas hasta la fecha en la experiencia humana. La validación explícita en la vida real es relativamente infrecuente. Más raro aún es oírla anunciada en público, lo cual es la forma de aprobación más potente porque transmite nuestro valor a toda la comunidad. ¿Cuándo fue la última vez que cincuenta, sesenta o setenta personas te aplaudieron en público en el mundo físico? ¿Quizás hace algunos años, si es que eso te ha ocurrido alguna vez? Pues en las redes sociales, eso ocurre una mañana cualquiera.

Además, las plataformas le añadieron un toque poderoso: un contador en la parte inferior de cada publicación indicando el número de «me gusta», retuits o votos favorables que había recibido: una cuantificación permanente de la aprobación social de todas y cada una de las publicaciones. Así fue como incluso LinkedIn, un tablón de

anuncios que alberga *curriculum vitae*, se convirtió en un sitio web para hacer contactos y fue vendido a Microsoft por 26.200 millones de dólares. Había añadido unos distintivos a los perfiles de los usuarios que indicaban el tamaño de su red. «Aunque en ese momento no podías hacer nada útil con LinkedIn, ese simple icono tuvo un poderoso efecto, ya que sacaba partido del deseo de las personas de no quedar como unos perdedores»,²⁵ ha dicho B. J. Fogg, director del Persuasive Tech Lab, de la Universidad de Stanford. Pero en 2020, incluso el cofundador de Twitter y luego director ejecutivo de la empresa, Jack Dorsey, admitió que había llegado a dudar del razonamiento que había desembocado en el botón de «me gusta» y sobre todo «en el hecho de que el botón tuviese un número asociado».²⁶ Pese a que no se comprometió a eliminar la función, reconoció que había generado «un incentivo que puede ser peligroso».

De hecho, el incentivo es tan poderoso que aparece incluso en escáneres cerebrales. Cuando recibimos un «me gusta», se enciende la actividad neuronal en una parte del cerebro llamada núcleo *accumbens*: el área que activa la dopamina.²⁷ Los sujetos con un núcleo *accumbens* más pequeño —un rasgo asociado con las tendencias adictivas— pasan periodos más largos en Facebook.²⁸ Y cuando las personas que pasan mucho rato en Facebook reciben un «me gusta», esa materia gris presenta mayor actividad que en usuarios que pasan menos tiempo, al igual que los adictos al juego que han sido condicionados para exaltarse cada vez que tiran de la palanca.

Pearlman, la empleada de Facebook que había contribuido a introducir el botón de «me gusta», descubrió eso tras dejar Silicon Valley, en 2011, para dibujar cómics. Publicitaba su obra, ¿cómo no?, en Facebook. De entrada, sus cómics tenían éxito. Retrataban temas inspiradores relacionados con la gratitud y la compasión, y a principios de la década de 2010 los sistemas de Facebook fomentaban esas emociones. Hasta que en 2015 Facebook modificó sus sistemas para desfavorecer el *clickbait* o «ciberanzuelo», que tanto atrae la curiosidad, lo cual tuvo el efecto secundario de eliminar el apoyo artificial que la plataforma antes daba al contenido de Pearlman, de una emotividad agradable.

«Cuando Facebook modificó su algoritmo, mis “me gusta” cayeron en picado y me daba la sensación de que me faltaba el aire —dijo Pearlman en declaraciones a *Vice News*—. Así que, aunque podía

culpar de ello al algoritmo, algo en mi interior me decía: “No les gusto, no soy lo bastante buena”.»²⁹ La misma empresa en la que había trabajado había puesto el núcleo *accumbens* de su cerebro en su contra y le había despertado un deseo interno de recibir «me gusta» tan fuerte que eso tenía más peso que su buen juicio. Luego, igual que Skinner jugueteando con un sujeto de investigación, simplemente cerró el grifo de las recompensas. «De repente estaba comprando anuncios, solo para recuperar aquella atención», admitió.

Para la mayoría de nosotros, el proceso es más sutil. En lugar de comprar anuncios de Facebook, modificamos nuestras publicaciones y comentarios diarios para mantener el flujo de dopamina, por lo general sin darnos cuenta de que lo hacemos. Este es el verdadero «bucle de retroalimentación por validación social», como lo ha denominado Sean Parker: perseguir de forma inconsciente la aprobación de un sistema automatizado diseñado para poner nuestras necesidades en nuestra contra.

«Es muy habitual que los seres humanos desarrollemos cosas con la mejor de las intenciones y que esas cosas tengan consecuencias negativas no intencionadas»,³⁰ dijo a *The Guardian* Justin Rosenstein, exingeniero de Facebook que también había trabajado en el botón de «me gusta». «Si solo le damos importancia a la maximización de los beneficios, descenderemos en poco tiempo en una distopía —alertó Rosenstein—. Un motivo por el que pienso que es especialmente importante que hablemos sobre este tema ahora es que podríamos ser la última generación que pueda recordar cómo era la vida antes.»

LA HONDA

A pesar de la importancia que la atención y la aprobación ejercen sobre usuarios como Pearlman y del tirón adictivo de unos iconos que parecen salidos de un casino, la fuerza más poderosa de las redes sociales es la identidad. Es el estímulo que mejor funciona en los sistemas de esa tecnología y que, por tanto, dichos sistemas están diseñados para activar y generar por encima de cualquier otra cosa. Expresar la identidad, refinarla, ver y definir el mundo a través de su óptica. Este efecto renovó el funcionamiento de las redes sociales, ya que sus supervisores y sistemas automatizados evolucionaron hasta

centrarse por completo en la identidad, que era lo más útil para sus objetivos.

Para entender el poder de la identidad, empecemos por preguntarnos: ¿qué palabras describen mejor mi identidad? Puede que nos vengan a la cabeza la nacionalidad, la raza o la religión. Quizás nuestra ciudad, profesión o género. Nuestra percepción de quiénes somos deriva en buena medida del hecho de formar parte de grupos. Pero esta compulsión —sus orígenes, sus efectos en nuestra mente y nuestras acciones— «sigue siendo un profundo misterio para los psicólogos sociales»,³¹ escribió en 1979 Henri Tajfel, cuando se propuso resolverlo.

Tajfel había experimentado el poder de la identidad de primera mano. En 1939 Alemania ocupó su país natal, Polonia, mientras él estaba estudiando en París. Judío y temeroso por su familia, se hizo pasar por francés para alistarse al ejército de este país. Siguió con el engaño cuando fue capturado por los soldados alemanes. Después de la guerra, al conocer que su familia había sido exterminada, obtuvo la nacionalidad francesa y luego la británica. Esas identidades eran meras construcciones sociales: ¿cómo, si no, podía cambiar de una a otra como el que se pone un jersey distinto? Y, aun así, tenían el poder de provocar una actitud asesina o misericorde en otras personas de su entorno y habían arrastrado a todo un continente a la autodestrucción.

Las preguntas que aquello planteaba perseguían y fascinaban a Tajfel. Él y varios compañeros suyos iniciaron el estudio de ese fenómeno, al que denominaron teoría de la identidad social.³² Rastrearon sus orígenes hasta un desafío formativo de los inicios de la existencia humana. Muchos primates viven en grupos reducidos. Los humanos, por el contrario, surgieron en grandes colectivos, donde el parentesco no era suficiente para unir a miembros del grupo que en su mayoría no estaban relacionados. El dilema era que el grupo no podía sobrevivir sin que cada miembro contribuyera al conjunto; y ningún individuo, a su vez, podía sobrevivir sin el apoyo del grupo.

Tajfel demostró que la identidad social es la manera de establecer vínculos con los miembros del grupo y ellos con nosotros.³³ Es el motivo por el que nos sentimos obligados a colgar una bandera en el balcón, a ponernos una camiseta de nuestra universidad o a pegar un

adhesivo en el parachoques del coche. Eso le dice al grupo que valoramos nuestra afiliación como una extensión de nosotros mismos y que, por tanto, se puede confiar en que serviremos al bien común.

El impulso que nos lleva a cultivar una identidad compartida es tan fuerte que creamos una incluso de la nada. En un experimento,³⁴ los investigadores asignaron a los voluntarios una etiqueta —de dos posibles— simplemente echando una moneda al aire, luego les hicieron jugar a un juego. Todos los participantes se mostraron más generosos con las personas que llevaban la misma etiqueta que ellos, aun sabiendo que la división no tenía ningún significado. Esa misma conducta se ha detectado en muchos experimentos y situaciones reales,³⁵ en las que los individuos aprovechan de forma sistemática cualquier excusa para separar entre un «nosotros» y un «ellos» y para expresar desconfianza, o incluso hostilidad, respecto a los miembros del exogrupo. Durante las pausas para comer en el set de la película de 1968 *El planeta de los simios*, por ejemplo, los extras se separaban de un modo espontáneo en mesas distintas según si hacían de chimpancé o de gorila.³⁶ Durante años, Charlton Heston, estrella del filme, contaba que ese «instinto de segregación» le había parecido «de lo más espeluznante». Cuando se filmó la secuela, un grupo distinto de extras repitió exactamente la misma conducta.

Los prejuicios y la hostilidad siempre han motivado ese instinto. Las tribus de cazadores-recolectores a veces competían por recursos o territorio. La supervivencia de un grupo podía requerir la derrota de otro. A raíz de eso, los instintos de la identidad social nos llevan a desconfiar y, si hace falta, a unirnos en contra de miembros del exogrupo.³⁷ Nuestra mente fuerza esos comportamientos provocando dos emociones en concreto: el miedo y el odio. Los dos son más sociales de lo que uno podría pensar. El miedo a una amenaza física del exterior nos hace tener un mayor sentimiento de camaradería con nuestro endogrupo, como si nos acercásemos enseguida a la tribu para proteger nuestra seguridad. También nos hace desconfiar más de las personas que percibimos como diferentes y nos vuelve más propensos a hacerles daño.³⁸ Pensad en la respuesta a los atentados del 11-S: hubo una ola de fervor patriótico y de empatía por el dolor de los conciudadanos, pero a eso le siguió un aumento brusco de los crímenes de odio contra musulmanes.

Estos instintos son profundamente sociales, así que las redes, al convertir cada clic en un acto social, sin duda los hacen aflorar. Y como las plataformas promueven cualquier sentimiento que genere más participación, a menudo producen esos instintos en su forma más extrema. El resultado puede ser una realidad artificial en la que el endogrupo siempre es virtuoso pero a la vez está asediado, el exogrupo siempre es una amenaza aterradora y casi todo lo que sucede es cuestión de nosotros contra ellos.

Al principio, la gratificación de la identidad por parte de las redes sociales no era dañina de una forma obvia, pero siempre había sido bien conocida. En 2012, un activista de izquierdas recaudó dinero de los cofundadores de Facebook y Reddit para poner en marcha Upworthy, que producía contenido pensado especialmente para difundirlo en las redes sociales. Comprobando de forma permanente qué lograba mayor difusión, Upworthy obtuvo por ingeniería inversa una fórmula para conseguir que algo fuera viral. Tenían éxito las listas numeradas. Y también los titulares que despiertan la curiosidad, que pedían que los usuarios los clicasen: «Nunca dirías qué dijo este entrenador para animar a sus jugadores». Pero una fórmula demostró tener una especial eficacia: los titulares que prometían retratar al endogrupo implícito del usuario (por lo general, los liberales) humillando a un despreciado exogrupo (los creacionistas, los racistas o las grandes empresas): «Un hombre responde a la pregunta de un intolerante con un zasca y se lleva una ovación del público».

Mientras tanto, una gran cantidad de periódicos recortaban la plantilla o cerraban porque internet había arrasado su modelo de negocio. Upworthy, con un presupuesto mínimo, había conseguido un público varias veces superior al de cualquier periódico. El sector, desesperado, tomó nota. Aparecieron o se reestructuraron organizaciones enteras en torno a la idea de conseguir ser virales. *BuzzFeed* se convirtió en un gigante de internet gracias a los artículos basados en listas que satisfacían el deseo de afirmación de la identidad social de los usuarios: «28 señales de que te criaron unos padres irlandeses» o «31 cosas que solo entenderás si eres de pueblo».

En 2014, yo fui uno de los periodistas del *Washington Post* que fundaron Vox, un portal de noticias que tenía la intención de aprovechar las posibilidades de internet. Nunca concebimos nuestro periodismo para que satisficiera los algoritmos de las redes sociales —

al menos, no de forma consciente—, pero los titulares estaban pensados con esos aspectos en mente. La estrategia más eficaz, aunque retrospectivamente tal vez deberíamos haber sido más precavidos al emplearla, era la del conflicto identitario. Liberales contra conservadores. La superioridad moral del antirracismo. La vergüenza de las permisivas leyes sobre el control de las armas. «La identidad era la honda —escribió Ezra Klein, fundador de *Vox*, en referencia a los medios digitales en un libro sobre la polarización—. Pocos se dieron cuenta, al principio, de que la manera de ganar la guerra de la atención era sacar provecho del poder de la comunidad para crear identidad. Pero los ganadores se impusieron con rapidez, a menudo utilizando técnicas cuyos mecanismos no comprendían del todo.»³⁹

Muchas veces, eso supuso la aparición de provocadores ultrapartidistas, granjas de clics con ánimo de lucro o auténticos estafadores. Sin respeto alguno por la justicia, la precisión o el bien común, llegaban a un público descomunal aprovechando o provocando conflictos identitarios. Quizás al principio las consecuencias no parecieran extenderse mucho más allá de internet. Pero habían llegado señales de alarma de lo más espeluznantes desde hacía años —y ahora vemos que eran de una claridad meridiana— de una zona del mundo donde había mucho en juego y a la que apenas se prestaba atención.

LA AVENTURA DE VUESTRA VIDA

¿Qué ocurre cuando toda una sociedad descubre internet de golpe, pasando de la noche a la mañana de una vida sin redes sociales a una vida dominada por estas plataformas? Aunque pueda sonar a algo imposible, este experimento ocurrió. Y fue en Birmania.

«Estoy convencido de que todos estáis a punto de vivir la aventura de vuestra vida —dijo Eric Schmidt, director ejecutivo de Google durante años, a un aula llena de estudiantes en el marco de una visita que hizo a este país del sureste asiático a principios de 2013—. Después de esta aventura con internet no habrá vuelta atrás.»⁴⁰

Durante décadas, este país —con la extensión del estado de Texas— de bosques tropicales, deltas fluviales repletos de arrozales y con un litoral bañado por el océano Índico, había vivido uno de los

aislamientos más herméticos del planeta. Una junta militar paranoica imponía la prohibición casi total de internet, los móviles, los medios extranjeros y los viajes internacionales. La tortura y la violencia represiva se ejecutaban con la peor combinación de incompetencia y crueldad. En 2011, el anciano dirigente del país fue sustituido por otro general de rostro adusto, Thein Sein, pero Sein resultó tener tendencias reformistas. Instó a los exiliados a regresar al país, relajó las restricciones a los medios de comunicación y liberó prisioneros políticos. Se distanció de China, el vecino del norte de Birmania, cada vez más arrogante, e inició negociaciones con Estados Unidos. Se levantaron las sanciones y se convocaron elecciones; en 2012, Barack Obama fue el primer presidente estadounidense en ejercicio que visitó el país.

Un actor secundario pero muy visible en la apertura del país, visto con buenos ojos tanto por los dirigentes birmanos como los estadounidenses, fue Silicon Valley. Prometieron que introducir el país a internet en poco tiempo modernizaría su economía y empoderaría a sus cincuenta millones de habitantes, lo cual consolidaría de manera eficaz la transición democrática. Varios meses después de la visita de Obama, Schmidt, actuando como embajador plenipotenciario de Silicon Valley, aterrizó en Rangún, la capital histórica de Birmania, para anunciar la llegada de las grandes tecnológicas. Flanqueado por el embajador estadounidense, dijo al público estudiantil: «Internet, una vez implantado, garantizará que la comunicación y el empoderamiento sean ley y costumbre de vuestro país».

Los dirigentes de Birmania también creían en el proyecto de futuro de Silicon Valley. Un periódico gestionado por el Estado advirtió a los ciudadanos de que «una persona sin una identidad en Facebook es una persona sin domicilio».⁴¹ El país se conectó a internet de una forma casi instantánea. Entre 2012 y 2015, los índices de adopción de internet se dispararon de un 0,5 % a un 40 %, en su mayor parte gracias a los teléfonos móviles baratos.⁴² Los precios de las tarjetas SIM se redujeron de 1.500 a 1,50 dólares.

Facebook desempeñó un papel destacado. Mediante acuerdos con empresas locales, consiguió que los móviles viniesen con una versión básica de la aplicación de Facebook ya instalada. En países pobres como Birmania, donde los ingresos medios son de unos tres dólares al día, los datos móviles pueden tener un precio prohibitivo. Para

superar este obstáculo y ganar la carrera para hacerse con los dos o tres mil millones de clientes más pobres del mundo, Facebook y otras compañías tecnológicas estadounidenses iniciaron la práctica del *zero-rating* o «programas de datos gratuitos», es decir, subvencionar a toda la población llegando a acuerdos con las compañías de telefonía locales para que no cobrasen nada por los datos consumidos al usar las aplicaciones de esas empresas. Birmania fue uno de los lugares donde se probó por primera vez y, para Facebook, fue un éxito abrumador. Un enorme porcentaje de la población aprendió a mandar mensajes y a navegar por internet exclusivamente a través de Facebook, hasta el punto de que muchas personas aún no son conscientes de que hay otras formas de comunicarse o de leer noticias por internet.⁴³

Llegué por primera vez a Birmania a principios de 2014. Aterricé en Rangún para informar sobre la incierta transición democrática del país. Me dio la sensación de estar en un sitio que se había quedado congelado a comienzos de los años sesenta, cuando los dirigentes militares lo habían aislado del mundo exterior. Los cortes de luz eran habituales y era infrecuente encontrar tecnología moderna; se acababan de instalar los primeros cajeros automáticos internacionales. El centro de la ciudad aún estaba dominado por las oficinas coloniales británicas, medio derruidas y conquistadas por la hiedra. Muchas calles del centro no estaban asfaltadas todavía y, en las primeras horas del día, eran ocupadas por cientos de monjes descalzos. Los sacerdotes, con túnicas de color naranja o carmesí, reverenciados por la mayoría budista, muy devota, son omnipresentes en Birmania.

Al entrevistarme con políticos y activistas, vi el futuro de Birmania más inestable de como se había pintado. El ejército todavía controlaba unos vestigios del poder que parecía reacio a ceder. Entre el clero, estaba en auge un sector extremista. Y las redes sociales, disponibles desde hacía poco, estaban repletas de racismo y conspiraciones. En el mundo virtual, por todas partes circulaban comentarios airados sobre minorías traicioneras.

En mis conversaciones se mencionaba una y otra vez un nombre preocupante: Wirathu. Este monje budista había sido encarcelado por sus sermones vilipendiosos durante la década anterior y acababan de liberarlo a raíz de una amnistía general. Enseguida se había creado perfiles en Facebook y YouTube. Ahora, en lugar de desplazarse de templo en templo por todo el país esparciendo odio, utilizaba las

plataformas para llegar a buena parte del territorio, incluso varias veces al día. Acusaba a la minoría musulmana del país de unos crímenes horribles, y lo hacía mezclando rumores con invenciones descaradas. Sobre todo en Facebook, sus publicaciones se propagaban entre unos usuarios que las tomaban como verdades,⁴⁴ y eso creaba una realidad alternativa determinada por la conspiración y la rabia, lo que catapultó a Wirathu hasta un nuevo grado de popularidad.

Una investigadora de Stanford que había trabajado en Birmania, Aela Callan, se reunió con altos directivos de Facebook a finales de 2013 para alertarlos de que los discursos de odio estaban infestando la plataforma, según le contó más adelante al periodista Timothy McLaughlin. Para un país con cientos de miles de usuarios —que pronto serían millones—, Facebook solo tenía contratado un moderador para revisar contenido en birmano, la lengua mayoritaria del país, lo cual dejaba la plataforma sin una vigilancia eficaz. Los directivos dijeron a Callan que Facebook seguiría adelante de todos modos con su expansión en Birmania.⁴⁵

A principios de 2014, Callan alertó de nuevo a Facebook: el problema estaba empeorando, y con él la amenaza de la violencia. Una vez más, hubo pocos cambios. Pasados unos meses, Wirathu compartió una publicación en la que afirmaba falsamente que dos musulmanes que regentaban sendas tiendas de té en la ciudad de Mandalay habían violado a una mujer budista. Publicó los nombres de los dos hombres y de sus establecimientos, y calificó la agresión ficticia como el acto que daría comienzo a un levantamiento musulmán masivo contra los budistas. Instó al Gobierno a organizar redadas en las viviendas de los musulmanes y las mezquitas a modo de ataque preventivo: una demanda habitual de los genocidas, cuyo mensaje implícito es que los ciudadanos de a pie deben hacer aquello que las autoridades no hacen. La publicación se hizo viral y dominó los canales de noticias de todo el país. Muchos usuarios indignados se unieron en torno a aquel odio visceral y se animaban unos a otros a aniquilar a sus vecinos musulmanes. Hubo cientos de disturbios en Mandalay, en los que se atacaron negocios y propietarios musulmanes. Hubo dos fallecidos y muchos más heridos.

Al extenderse los disturbios, un alto funcionario del Gobierno llamó a alguien que conocía en las oficinas de la consultora Deloitte en Birmania para pedirle ayuda: quería ponerse en contacto con

Facebook. Pero ninguno de los dos logró hablar con nadie de la compañía. Desesperado, el Gobierno bloqueó el acceso a Facebook en Mandalay. Los disturbios se enfriaron. Al día siguiente, empleados de Facebook respondieron por fin al representante de Deloitte, no para interesarse por los episodios violentos, sino para pedirle explicaciones por el bloqueo de la plataforma. En una reunión celebrada dos semanas después con el funcionario y otras personas, un representante de Facebook dijo que estaban trabajando para mejorar su capacidad de respuesta ante contenidos peligrosos en Birmania. Pero, si la compañía hizo algún cambio, el efecto fue indetectable en la plataforma. En cuanto el Gobierno levantó el bloqueo virtual, los discursos de odio, y la audiencia de Wirathu, no hicieron más que crecer. «A partir de ese incidente en Mandalay, si no antes, Facebook lo sabía —le dijo David Madden, un australiano que dirigía el acelerador de empresas emergentes más grande de Birmania, a McLaughlin—. No es que se vea solo ahora, a toro pasado. El alcance de ese problema era notable y ya se veía con claridad.»

Sin capacidad o sin voluntad para plantearse que su producto podía ser peligroso, Facebook siguió expandiéndose por Birmania y otros países en vías de desarrollo y con niveles de control insuficientes. Fundamentaba esa decisión por completo en un credo de Silicon Valley encaminado al autoenriquecimiento que Schmidt había recitado en aquella primera visita en Rangún: «La manera de combatir los discursos negativos es que haya más discursos. Más comunicación, más voces».

Todo es como el Gamergate

UNA NUEVA ERA

Era agosto de 2014, un mes después de los disturbios de Mandalay. Mientras Zoë Quinn estaba brindando con unos amigos en un bar de San Francisco para celebrar su vigesimosexto cumpleaños, la web social se le echó encima con una fuerza tal que alteró la trayectoria de internet y de casi todo lo que ha aparecido desde entonces. Su móvil hizo un zumbido para notificarle la llegada de un mensaje de texto de un amigo: «Han colgado info tuya muy *heavy*».¹ Un programador llamado Eron Gjoni había publicado en su blog un relato inconexo de 10.000 palabras sobre la breve relación y la ruptura entre ambos, con pantallazos de correos electrónicos, mensajes de texto y mensajes privados de Facebook.

Quinn, una programadora de videojuegos con el pelo teñido con Technicolor, era una figura habitual en foros y plataformas sociales *geek*. Había recibido cierta atención de los críticos por obras de arte *indie* como *Depression Quest*, un simulador basado en texto sobre la depresión clínica y por su feminismo sin pelos en la lengua. Hacía publicaciones con frecuencia, a veces con un tono llamativo, a favor de alguna causa que estaba ganando impulso entre creadores de juegos y periodistas con ideas parecidas a las suyas: ampliar el gusto por los videojuegos y la cultura de los fans más allá de su territorio tradicional de los chavales jóvenes *geeks*. Pero algunos círculos virtuales de jugadores estaban furiosos al ver a transgresoras feministas que intentaban, a sus ojos, corromper el *hobby* que se había convertido, en un mundo que pintaba a muchos de los primeros que se obsesionaron con las redes sociales como unos individuos hostiles y confusos, en una suerte de espacio seguro. Ese debate iba más allá de si los juegos de príncipe-salva-a-princesa-tetuda podían dejar sitio a productos no convencionales como los de Quinn o a las chicas *gamers*; el debate giraba en torno a una identidad masculina *geek*, cuyos

miembros se veían atacados. Se sentían identificados con el relato de rabia y resentimiento de Gjoni.

Un detalle concreto de la publicación de Gjoni había dado tumbos por foros de videojuegos, había terminado en plataformas mayoritarias, había llegado a los amigos de Quinn y había vuelto al teléfono móvil de la chica. Gjoni afirmaba que Quinn se había acostado con un crítico de videojuegos a cambio de una reseña positiva de Depression Quest. Su acusación podía desmentirse con facilidad: la supuesta reseña ilícita ni siquiera existía. Pero la verdad apenas importaba. Los usuarios de las subsecciones sobre videojuegos del tablón de mensajes 4chan, un centro de la cultura *nerd*, y sobre todo los de Reddit, un extenso portal de debate que se había convertido en una megalópolis en ebullición en el corazón de la web social, tomaron la afirmación de Gjoni como una justificación de su desconfianza, lo cual consolidó el relato de millones de usuarios de las plataformas.

La publicación de Gjoni también se interpretó como algo que promovía la justicia sin miramientos que a menudo se lleva a cabo en la web social: el acoso colectivo. Hasta el punto de que tiempo después un juez prohibió a Gjoni escribir nada más acerca de Quinn. Y desde luego, si el acoso era el objetivo de Gjoni, su publicación consiguió el efecto deseado. «solo quiero ver a zoe recibir su merecido»,² escribió un usuario de 4chan en un chat en el que se organizaba la iniciativa destinada a, como describía otro, «hacerle la vida irremediablemente imposible», incluso «acosándola para que se suicide».

«Intenté centrarme en la conversación del bar —escribió más tarde Quinn—, pero no podía desviar la atención de las vibraciones que hacía mi móvil. Era como contar los segundos entre los truenos para ver lo lejos que está la tormenta y saber que se está acercando.»³

Se escabulló de la fiesta y se fue a casa para rastrear de dónde salía aquel odio virtual. Para entonces ya había cientos de mensajes que la animaban a suicidarse y que prometían atormentar a su familia si no lo hacía, amenazas creíbles por ir acompañadas de listas de direcciones y números de teléfono de parientes suyos. En algunos casos se difundían fotos tuyas retocadas con contenido pornográfico. En otros se publicaban detalles personales como su número de la Seguridad Social, con lo cual demostraban su poder de inmiscuirse en

su vida. «Si alguna vez me entero de que te han invitado a hablar en un acto al que yo asista, literalmente te voy a matar —escribía uno—. Eres peor que la mierda y mereces que te pegue, te mutile, te mate y, además, adorne tu cadáver podrido meándome mil veces encima de él.»⁴

La rabia contra Quinn y la presuntamente corrupta prensa de los videojuegos invadió buena parte de 4chan y Reddit, y luego de YouTube. En las tres plataformas, unas comunidades enormes se obsesionaron cada vez más con ese escándalo inventado al que llamaron Gamergate. Pero lo que había empezado como un episodio más de troleo virtual —aunque sus dimensiones fueran inusuales— se convirtió de forma paulatina en algo más, algo nuevo. El Gamergate alteró más cosas aparte de la vida de aquellos que fueron atacados. Hizo caer estrepitosamente los extremos de la web social encima de la vida estadounidense mayoritaria, lo cual acabó para siempre con la separación entre el espacio digital y el no digital, entre la cultura de internet y la cultura a secas.

También dio comienzo a un nuevo tipo de política, definido por los rasgos fundacionales de las redes sociales: una cultura digital construida en torno a unos hombres jóvenes nihilistas, a los sueños de revolución destructiva de Silicon Valley y a unas plataformas diseñadas de maneras que sobrecargan la identidad para convertirla en una cuestión de conflicto existencial y totalizante. Otras comunidades, ya fuesen tan reducidas como las madres *hippies* ecologistas de Facebook o tan grandes como la derecha política estadounidense, ya estaban siguiendo una trayectoria de un antagonismo parecido. El Gamergate en sí no fue más que una versión de eso, pero una versión que trajo consigo las semillas de otras que vendrían. Los extremistas «íncels» violentos, una extrema derecha rejuvenecida y reinventada, y su rama de la *alt-right* o derecha alternativa, adaptada a los jóvenes: todos estos movimientos aprovecharon el Gamergate, al igual que una cepa transformadora de otro movimiento que empezaba a cobrar forma: el trumpismo. Entre los analistas y los periodistas que trataban de entender esa nueva era, en la que las normas que gobernaban la web social empezaron a gobernarnos a todos, se hizo popular una frase: «Todo es como el Gamergate».

Pasaría otro año hasta que la mayor parte de las personas ajenas

al revuelo inmediato del Gamergate —que se propagó a gran velocidad por los medios de comunicación, el mundo del entretenimiento y casi cualquier comunidad virtual— percibieron las consecuencias. Pero al principio se trataba de una venganza, de unos hombres y chavales que vivían en internet que se sentían abandonados por la sociedad estadounidense y exigían venganza contra quienquiera que las redes sociales, donde ellos encontraban refugio, les hubieran enseñado a culpar.

«Si piensas que tu enemigo es un símbolo y no una persona, de repente hay un montón de putadas inhumanas que tienes el ancho de banda emocional para hacer —dijo más adelante Quinn a un periodista—. Y lo sé porque yo he sido una cabrona. Si el Gamergate le hubiese ocurrido a otra persona, años antes, yo probablemente habría estado en el lado malo. ¿Que era una adolescente de mierda con una enfermedad mental que tenía una vena misógina y a la que le encantaban los videojuegos? Pues sí.»⁵

A lo largo de los siguientes meses, los *gamergaters* atacaron a muchas personas que levantaron la voz en defensa de Quinn, que criticaron sus métodos o la cultura de los videojuegos por internet, o que corría el rumor de que estaban involucrados en la conspiración de la que los propios *gamergaters* se habían convencido unos a otros. Se volvieron en contra de escritoras y periodistas de mentalidad parecida y causaron unos estragos que ni siquiera los críticos más lúcidos de las redes sociales creían posible. Uno de los métodos favoritos era el *swatting*:⁶ llamar al teléfono para emergencias 911 contando una historia inventada según la cual hay una toma de rehenes en el domicilio de la persona que se quiere incordiar para desencadenar, a menudo con éxito, una redada de un equipo de las fuerzas especiales SWAT que, en medio de la confusión, pueda acabar con que la policía dispare a la persona en cuestión o a su familia. (En 2017, la policía detuvo a un hombre que llamó y desencadenó una actuación fatal de los SWAT a raíz de una disputa sobre videojuegos por internet.⁷ Fue condenado a veinte años de cárcel en el marco de un acuerdo judicial por una amenaza de bomba que había anunciado. También han sido detenidos otros dos *swatters* en serie,⁸ cada uno responsable de decenas de avisos falsos, en su mayor parte por peleas sobre videojuegos por internet. Sin embargo, por lo general los acosadores virtuales no suelen afrontar consecuencias penales. Aunque las autoridades

judiciales han empezado a tomarse más en serio esas actividades, investigarlas requiere mucho tiempo y, como son tan descentralizadas, es casi imposible impedir las o disuadir a sus autores. Cuando la policía logra descubrir la identidad de un acosador —y no digamos actuar—, por lo general el acoso ya ha terminado. Y procesar a uno o dos acosadores especialmente agresivos, aunque pueda ser repositivo para la víctima, no sirve de gran cosa para impedir que otros cien usuarios reproduzcan el mismo efecto.) Las mujeres que trabajaban en las empresas de videojuegos, puesto que eran vistas como cómplices, también fueron objeto de ataques. Algunas, temerosas de que sus hijos pudiesen correr peligro, abandonaron el sector por completo.

Cuando un portal de noticias sobre el sector acusó a los *gamergaters* de dejar la cultura de los videojuegos en una situación «un poco embarazosa»,⁹ los usuarios inundaron el principal anunciante del portal (Intel) de quejas inventadas, con lo cual durante un tiempo convencieron al gigante empresarial de que retirase su publicidad. Y lo mismo hicieron con *Gawker*. Los periodistas que trabajaban en lo que hasta entonces había sido un mundo de los videojuegos de ritmo tranquilo aprendieron a eliminar información personal de internet y, cuando les llegaba una atención no deseada, alertaban a los miembros de su familia de la posibilidad de recibir amenazas de muerte o ciberataques.

«Lo que realmente me pareció extremo fue cuando la ola que se convirtió en el Gamergate atacó a Samantha Allen, que escribía para Polygon —me contó más adelante Brianna Wu, una desarrolladora de videojuegos independiente—. Literalmente, la echaron de su trabajo. Así que no era un drama virtual. Había mujeres a quienes les destruían la carrera profesional.»

El acento del Misisipi de Wu era una constante en los debates y los pódcast del sector, donde, mucho antes del Gamergate, Wu había reclamado que se reformaran las tendencias más infantiloides del sector. Cuando el Gamergate atacó a sus amigas, se sintió obligada a levantar la voz. Ya preveía que sería objeto de contraataques, pero no era consciente de la ferocidad de la respuesta. Entre finales de 2014 y comienzos de 2015, miles de personas le mandaron correos basura con imágenes gráficas e insultos, lo que dejó inservibles las plataformas que usaba para publicitar su trabajo. Las amenazas de muerte venían acompañadas de planos de su casa y fotos de su familia. Un amigo la

alertó de que había foros en los que los usuarios se animaban unos a otros a llevar a la práctica las amenazas.

«Estoy segura de que parte de eso era solo una táctica intimidatoria», dijo Wu. Pero hubo una amenaza en concreto que le hizo pensar que algunas podían ir en serio. Incluso años después, recordaba el mensaje palabra por palabra: «Hola, puta, ¿sabes qué? Sé dónde vivís Frank y tú. Si tenéis hijos, también van a morir. No has hecho nada de provecho con tu vida. Voy a cortarle a tu marido su diminuto pene asiático y te voy a violar con él hasta que sangres». Cuando lo puso en conocimiento de las autoridades, le aconsejaron que se fuera de su casa. Pasados varios meses, un *youtuber* colgó un vídeo en el que salía él mismo con una máscara de una calavera, blandiendo un cuchillo que prometía utilizar contra Wu. El incidente se incluyó en un episodio de *Ley y Orden: Unidad de Víctimas Especiales*. «Ice-T asesina a Logan Paul en una azotea para salvar al personaje que está inspirado en mí. Te lo juro por Dios, no podía creérmelo cuando lo oí», dijo con una carcajada. Las amenazas han continuado. Poco antes de conversar con ella, en junio de 2020, le habían tirado un ladrillo por la ventana que le había roto el cristal.

Cuando el Gamergate entró en la conciencia pública, Wu echó mano de sus contactos en las redes sociales y los presionó para que, por lo menos, limitasen las campañas de acoso que salían de sus sistemas. Pero los conocidos de Silicon Valley con quienes habló, en su mayor parte hombres jóvenes blancos, parecía que nunca se hubiesen planteado que el odio y el acoso pudieran tener consecuencias reales, y mucho menos habían pensado cómo contenerlas. «No es que sean malas personas —dijo Wu—. Sencillamente no han vivido ciertas experiencias que sí han vivido muchas mujeres, personas *queer* y personas de color.»

Las compañías menos sensibles fueron Facebook, que no se involucró en absoluto en su caso, y Reddit, uno de los sitios donde había empezado el Gamergate. Cuanto más interactuaba con los operadores de la plataforma o exploraba el veneno que emanaba de sus sitios web, más sospechaba Wu que existía un peligro más generalizado. «El *software* define cada vez más el mundo que nos rodea»,¹⁰ escribió a principios de 2015. Las plataformas y las aplicaciones «crean nuestras realidades sociales: cómo hacemos amigos, cómo conseguimos trabajo y cómo interactúa la humanidad».

Pero habían sido diseñadas con poca participación de personas externas a la estrecha cosmovisión o la demografía de Silicon Valley. «Esos sistemas son la próxima frontera de la evolución humana y entrañan un peligro cada vez mayor para nosotros —concluyó Wu, y agregó una opinión que en ese momento se consideraba exagerada—: Hay muchísimo en juego.»

Aquella transformación se había puesto en marcha cuarenta años antes, con una generación de fabricantes de ordenadores de Silicon Valley que se veían a sí mismos como unos revolucionarios destinados a demoler por completo el *statu quo* estadounidense y que construyeron las redes sociales, de forma muy explícita, como la herramienta con que lo harían. Pero su nueva sociedad digital, concebida como una futura sustitución de todo lo que había existido antes, no se diseñó tanto para alcanzar la liberación como para atizar la ira y los conflictos, lo que tuvo lugar gracias a un pecado original del capitalismo de Silicon Valley y, en los años noventa, un giro fatídico en la economía de los juguetes. El resultado fue un mundo digital que a principios de la década del año 2000 ya cursaba con una extraña mezcla de chovinismo masculino *geek* y, aunque al principio se descartó la idea, con el extremismo de derechas.

El Gamergate anunció el inicio de nuestra nueva era, en que la vida estadounidense estaría moldeada por los incentivos y normas de las redes sociales, desde unas plataformas situadas más allá de las afueras de la sociedad convencional. En pocos años, esas plataformas convertirían el Gamergate y sus ramificaciones en movimientos de alcance nacional, los llevarían a los hogares de millones de personas recién llegadas al mundo digital y las movilizarían para que apoyasen un movimiento que, muy pronto, se instalaría en la Casa Blanca.

LOS REVOLUCIONARIOS

La etapa en que nació la revolución de los ordenadores coincidió con un periodo de terribles convulsiones en la sociedad estadounidense. Asesinatos, disturbios, la derrota en Vietnam y la vergonzosa dimisión de Richard Nixon suscitaron una profunda hostilidad contra la autoridad centralizada, así como una contracultura cuyas extravagancias desorbitadas parecían apropiadas para los tiempos.

Pero, aunque más adelante los expertos en *marketing* reescribirían la era de los ordenadores como una época de soñadores iconoclastas, en realidad empezó con personas como Douglas Engelbart, un ingeniero de investigación naval. Durante los años sesenta, Engelbart, con financiación procedente de la NASA y el Pentágono, trabajó con empeño en una máquina que usaba semiconductores para almacenar y mostrar información.¹¹ Pero, a diferencia de los mastodontes que funcionaban con tarjetas perforadas al estilo de IBM, sería lo bastante fácil para que la utilizarasen personas no expertas.

El aparato que Engelbart presentó con orgullo en una demostración pública en 1968 contaba con la primera interfaz gráfica. También incluía el primer ratón de la historia. Incluso podía intercambiar información con otras máquinas a través de un módem. Su demostración desató una tormenta de entusiasmo en Silicon Valley, que veía el nacimiento de un sector totalmente nuevo. Los intelectuales públicos, imbuidos del entusiasmo contracultural del momento, anunciaron el aparato como un paso adelante en el desmantelamiento de las estructuras de poder y la construcción de una nueva sociedad desde abajo. *Future Shock*, un megasuperventas publicado en 1970, pronosticaba una «revolución tecnológica» que permitiría que los individuos tuviesen más poder que las instituciones. El sociólogo Ted Nelson, amigo de Engelbart, escribió *Computer Lib/Dream Machines* [Liberación informática/Máquinas soñadas], cuyo título («Lib» es la abreviatura de *liberación*) e imagen de cubierta — con un puño levantado — transmitían ya buena parte del mensaje.

La mitología caló enseguida en un sector dispuesto a redefinirse. En 1971, una revista de negocios acuñó el término «Silicon Valley» en referencia al negocio de los transistores de silicio iniciado diez años antes por William Shockley y sus discípulos. Al año siguiente, se publicó un entusiasta perfil de unas oficinas de investigación de una empresa de Silicon Valley nada más y nada menos que en la revista *Rolling Stone*. Presentaba a los ingenieros como unos excéntricos greñudos que nos traerían «libertad y rareza» mediante un producto — el ordenador — que a la mayoría de la población le parecía un aparato aburrido e intimidante que estaba en su despacho o su universidad.

Los ingenieros de todo Silicon Valley interiorizaron gratamente aquellos halagos como si fueran un hecho. A partir de entonces, escribir código en un terminal ya no era desarrollo comercial de

producto, era «hackear». «Nosotros somos los verdaderos revolucionarios del mundo hoy en día, no los chavales pelilargos y barbudos que destrozaban los colegios hace unos años»,¹²le dijo Gordon Moore, uno de los cofundadores de Intel, a un periodista. Mientras que, tras la dimisión de Nixon, las excentricidades contraculturales disminuyeron en el resto de Estados Unidos, en Silicon Valley se mantuvieron, gracias en parte a la llegada, en 1974, del Altair 8800, el primer ordenador suficientemente pequeño y barato para utilizarse en los hogares. Las máquinas eran «abiertas», es decir, que cualquiera que tuviese los conocimientos necesarios podía modificar o sustituir componentes. Los ingenieros de Silicon Valley crearon clubes para hacer probaturas con ordenadores después del trabajo con nombres como Homebrew Computer Club (‘club de ordenadores caseros’) o People’s Computer Company (‘compañía popular de ordenadores’). En boletines informativos y reuniones regulares, codificaron su autoimagen revolucionaria y la convirtieron en algo semejante a una doctrina. El boletín informativo de PCC publicaba orientaciones técnicas junto con tratados sobre la futura utopía libertarista.¹³

Los encuentros del Homebrew originaron una generación de empresas emergentes, entre ellas Apple Computer. A medida que los ordenadores personales se expandieron más allá del nicho de mercado de los aficionados, Apple adelantó a sus competidores gracias a su tecnología y, sobre todo, a su *marketing*. Vendía la imagen de «libertad y rareza» a los *baby boomers* como algo que era nostálgico —recordaba a la contracultura de los años sesenta— y a la vez permitía quedar bien. En 1984, mientras el negocio se disparaba año tras año, Apple emitió un anuncio durante la Super Bowl de una mujer lanzando un martillo contra una pantalla de vídeo en la que aparecía un gobernante totalitario. El consumismo como revolución; un anarquismo de *hackers* mezclado con el capitalismo desacomplejado de los años ochenta con Reagan. «Es peligroso —dijo la historiadora Margaret O’Mara—, porque aquel mito se convierte en la realidad de Silicon Valley.»

Lo de derribar estructuras de poder era, de entrada, más que nada algo retórico. Pero algunos fanáticos dieron forma a aquellos ideales y los llevaron hasta un extremo casi mesiánico creando lo que llamaron el WELL, la primera red social con consecuencias reales. El

proyecto había surgido de una revista dirigida por un exsocio de Ken Kesey llamado Stewart Brand,¹⁴ que se había pasado la década de los sesenta conduciendo entre comunas *hippies* de California vendiendo herramientas que llevaba en su camioneta. Lo había denominado Whole Earth Truck Store ('tienda camioneta de la tierra entera'). Al establecerse en el valle de Santa Clara, lo convirtió, en 1968, en el Whole Earth Catalog ('catálogo de la tierra entera'). El nombre era un chiste: daba consejos a los lectores sobre cómo fabricar los productos por su cuenta, junto con artículos que promovían el comunismo *hippie*.

Los ejemplares eran omnipresentes en los inicios de Silicon Valley. Más adelante Steve Jobs lo calificó como «una de las biblias de mi generación». Brand, después de absorber las promesas de liberación de Silicon Valley, usó su revista y su credo *hippie* para repetírselo a sus lectores como una orden: solo vosotros podéis terminar lo que se inició en los sesenta. «Creo que los *hackers* —dijo Brand en un congreso del sector celebrado en 1984— son el grupo de intelectuales más interesante y eficaz desde los artífices de la Constitución de Estados Unidos.» Cuando una empresa de teleconferencias ofreció a Brand la idea de convertir su revista en un tablón de mensajes, él creó el Whole Earth 'Lectronic Link ('enlace electrónico de la tierra entera'), el WELL. Los amigos que tenía en las esferas más altas del sector se unieron todos, y les siguió buena parte de Silicon Valley, lo cual lo convirtió en un centro de actividad imprescindible.

Los fundadores del sitio lo concibieron como la plasmación en el mundo real de sus sueños de una utopía anarquista.¹⁵ Creían que una ausencia casi absoluta de normas conduciría a una comunidad autogobernada en la que las ideas se impondrían o se descartarían basándose en el mérito.¹⁶ En realidad, cuando entraron en tromba ingenieros cascarrabias y combativos, se impusieron las voces más vehementes y las opiniones más populares, pero, como los arquitectos del WELL representaban ambas cosas, se lo tomaron como la confirmación de que su meritocracia intelectual funcionaba. A partir de entonces, los arquitectos de la era de internet que se habían reunido por primera vez en el WELL considerarían siempre que el mayoritarismo sin matices era el ideal natural y lo incorporarían a todas las redes sociales que han ido llegando hasta la actualidad. Los

debates estridentes terminaron viéndose como la meritocracia más pura: si no lograbas vender tus proyectos o ganarte a la multitud, si te sentías acosado o rechazado, era porque tus ideas no habían triunfado en razón del mérito.

Los usuarios del WELL, esparcidos por todo el sector, fueron tejiendo la red de consumidores a imagen y semejanza de aquel proyecto: sin normas, sin gobierno, de uso gratuito y diseñado implícitamente para los hombres *geeks* que habían llenado el foro donde se fraguó. Y esa red no consistía solo en páginas web. Se trataba de una cibernsiedad que nos elevaba por encima de los anticuados procedimientos del mundo físico. «Rechazamos reyes, presidentes y votaciones. Creemos en el consenso aproximado y en la ejecución del código»,¹⁷ declaró David Clark, uno de los arquitectos de la web, en 1992.

En 1996, un antiguo miembro de la junta del WELL escribió el documento definitorio de la era web: «Declaración de independencia del ciberespacio».¹⁸ Dirigiéndose a los «Gobiernos del mundo industrializado», les anunciaba que «no tenéis soberanía allí donde nos reunimos nosotros». La web sería «una civilización de la Mente», gobernada por la voluntad colectiva de sus usuarios. Era una ideología que invadió en poco tiempo la cultura mayoritaria, representada en películas como *La red* o *Matrix*, que presentaban a los programadores como la nueva vanguardia contracultural, rebeldes de kung-fu que romperían las cadenas de la servidumbre humana.

El manifiesto consagraba un ideal en particular: la libertad de expresión total. Al igual que en el WELL, ese se convertiría en el mecanismo de la web para autogobernarse, su primer mandamiento y su mayor regalo para el mundo. Sus preceptos siguen siendo el texto fundacional del sector de las redes sociales. «A nuestro departamento jurídico y a nuestro director ejecutivo les gusta decir que somos el ala más favorable a la libertad de expresión del partido por la libertad de expresión»,¹⁹ había contado el máximo responsable de Twitter en el Reino Unido. Y Zuckerberg calificó la libertad de expresión como «el ideal fundacional de la compañía».²⁰

No obstante, lo que impulsa sobremanera la ideología subyacente de Silicon Valley es la ambición originaria de terminar haciendo una revolución de raíz, una profecía que las compañías de redes sociales se sienten destinadas a hacer realidad. Mientras que Apple era una

«empresa innovadora», Facebook era una «empresa revolucionaria», le dijo un Zuckerberg de veintidós años a un potencial trabajador.²¹ En una entrevista en televisión afirmó: «En cierto modo, estamos llevando a cabo una reconfiguración drástica desde abajo»,²² un compromiso que puso negro sobre blanco en una carta a los accionistas.²³ Zuckerberg creía que Facebook podía y debía hacer eso, porque estaba dirigida por ingenieros, cuya mirada pura los acompañaría a lo largo del todo el camino. «Hay esa cosa fundamental de que, cuando eres joven, te fijas en algo y piensas: eso puede ser mejor. Puedo desmontar ese sistema y mejorarlo —dijo una vez Zuckerberg en un viaje a Nigeria, donde prometió que Facebook ayudaría a avanzar a toda África—. De verdad pienso que esta es la mentalidad de los ingenieros. Tal vez sea más un conjunto de valores que una mentalidad.»²⁴ Pero el idealismo —la creencia en que las empresas emergentes que consiguieran más usuarios podrían y deberían rehacer toda la sociedad— reflejaba una soberbia que resultaría ser catastrófica. «El motivo por el que nosotros los *nerds* no encajábamos era que en ciertos aspectos íbamos un paso por delante»,²⁵ escribió una vez Paul Graham, el inversor cuya incubadora había impulsado Reddit. Los futuros habitantes de Silicon Valley eran «sistemáticamente impopulares» de niños, sostenía, porque «ya estábamos pensando en las cosas que importan en el mundo real, en lugar de dedicarnos a jugar a un juego riguroso pero casi siempre inútil como los demás».

Sin embargo, más importante aún que la creencia de Silicon Valley en una misión grandilocuente fueron los tipos de ingenieros que sus inversores encumbraron para liderar su revolución y a cuya imagen y semejanza se reconstruiría el mundo. «Todos parecen ser hombres *nerds* blancos que han dejado sus estudios en Harvard o Stanford y que no tienen vida social en absoluto», dijo una vez John Doerr, un legendario inversor tecnológico, refiriéndose a los creadores de empresas exitosos, y agregó que ese era el «patrón» en el que se basaba para seleccionar las empresas en las que invertía. Del mismo modo, Graham ha dicho que busca «*nerds*» e «idealistas» con «espíritu piratesco», a quienes «les guste saltarse las normas» y desafiar los actos de cortesía social.²⁶ «Esos tíos quieren hacerse ricos, pero quieren hacerlo cambiando el mundo.»²⁷ Peter Thiel, uno de los fundadores de PayPal y el primer inversor externo de Facebook, había animado a

fomentar a los contestatarios antisociales. «Parece que los individuos con una ineptitud social al estilo del síndrome de Asperger cuentan hoy en día con ventaja en Silicon Valley —escribió en un influyente libro sobre empresas emergentes—. Si eres menos sensible a los códigos sociales, es menos probable que hagas las mismas cosas que hacen todas las personas de tu alrededor.»²⁸ Los inversores consideraron que ese arquetipo representaba una forma extrema de meritocracia, basado exclusivamente en los resultados, tan puro que estaba por encima de insignificantes preocupaciones por la diversidad. En realidad, esos resultados, confirmados por el «patrón» de Doerr, reflejaban nada más y nada menos que una cultura tecnológica que era hostil a cualquier persona que no coincidiera con el ideal, tradicional en Silicon Valley, del hombre *geek* misántropo.

«No se da mucho valor a los actos de cortesía social —me dijo Margaret O'Mara—. Se tolera la rareza, en parte porque las personas raras tienen un historial demostrado. Esta es la otra dimensión de la cultura de Silicon Valley. Es como si todo el mundo fuera un capullo.» Este arquetipo de Silicon Valley derivó de las empresas que lo fundaron: Shockley Semiconductor Laboratory y el fabricante de componentes electrónicos Hewlett-Packard. A ambas las definía el desdén de sus fundadores cascarrabias por la cultura empresarial convencional y las estructuras de gestión, lo que los alejó de la Costa Este. Fomentaron una cultura laboral despiadadamente competitiva y fervientemente antijerárquica, que daba a los ingenieros rienda suelta y les imponía poca supervisión. Y contrataron a gruñones *antiestablishment* a su imagen y semejanza, con un historial demasiado informal y una personalidad demasiado difícil para empresas elegantes como IBM o Bell Labs. Al final, puesto que las personas beligerantes y con iniciativa que no habían terminado los estudios parecían estar por todas partes, Silicon Valley tomó esa personalidad como un signo de genialidad.

En la mayoría de los sectores, tales peculiaridades se atenuarían con el tiempo gracias a la llegada de nuevos perfiles y de generaciones posteriores. Pero, como ocurre a menudo en Silicon Valley, la fuerza oculta detrás de todo, que determina tanto la cultura como la economía, eran las inversiones de capital riesgo. La práctica de que los ingenieros se convirtieran en inversores de capital riesgo que elegían a la siguiente de generación de ingenieros dominantes preservó el

patrimonio genético ideológico dentro de unos límites incestuosos.

Aún hoy, Shockley está a tan solo cuatro o cinco pasos de casi cualquier personaje importante de las redes sociales. Una de las primeras personas a las que contrató, un ingeniero llamado Eugene Kleiner, cofundó más adelante Kleiner Perkins, la empresa de inversión que contrató a Doerr. Doerr, a su vez, invirtió en el capital inicial de Amazon y Google, donde sus consejos —lecciones que había aprendido de empleados de Shockley— se convirtieron en los cimientos del modelo de negocio de YouTube. Otro protegido de Doerr, el fundador de Netscape Marc Andreessen, fue uno de los principales inversores de Facebook y miembro del consejo de administración de la compañía, además de mentor personal de Mark Zuckerberg. Cofundó una empresa de capital riesgo que puso el dinero inicial para la creación de, entre otras empresas, Slack, Pinterest y Twitter.

Hay montones de interconexiones como esas, todas entre un reducido grupo de inversores y empresarios de opiniones parecidas. Yishan Wong, jefe de Reddit durante el Gamergate, se había labrado un nombre en PayPal, cuyos exalumnos guiaron buena parte de la era de las redes sociales. Uno de los primeros ejecutivos de PayPal, Reid Hoffman, utilizó los beneficios que obtuvo allí para crear LinkedIn y para ser uno de los primeros inversores de Facebook. Presentó a Zuckerberg a Thiel, quien se convirtió en el primer miembro del consejo de administración de Facebook.

Thiel, aprovechando también el éxito de PayPal, creó un fondo que hizo grandes inversiones en Airbnb, Lyft y Spotify. Durante todo ese tiempo, al igual que otros muchos inversores destacados, impuso sus ideales en las empresas que supervisaba. En los años noventa, fue coautor de un libro, *The Diversity Myth* [El mito de la diversidad], en el que se calificaba la inclusión deliberada de mujeres y minorías como una estafa que minaba el libre ejercicio intelectual. «Max Levchin, con quien creé PayPal, dice que las empresas emergentes deberían hacer que sus primeros empleados fuesen lo más parecidos posible en cuanto a personalidad —escribía Thiel—. En la empresa, todo el mundo debería ser distinto de la misma manera: una tribu de personas que piensan igual entregadas con devoción a la misión de la empresa.»²⁹

Ese, más que la raza o el género, era el rígido arquetipo en torno

al cual Silicon Valley diseñaba sus productos: hombres *geeks* blancos, misántropos, lógicos y despiadados. Durante la mayor parte de la historia del sector, esta predilección afectó a pocas personas más allá de las mujeres y las minorías que soportaban a duras penas su lugar de trabajo. Pero, con el inicio de la era de las redes sociales, el sector incorporó sus peores hábitos en empresas que luego introdujeron de tapadillo esos excesos —el chovinismo, una cultura del acoso, el mayoritarismo disfrazado de meritocracia— en los hogares y las mentes de miles de millones de consumidores.

EL TROLEO

La revolución llegó deprisa, propagada por la banda ancha hasta casi todas las ciudades y barrios. Trajo consigo, como se prometía, una libertad casi absoluta, por lo menos para quienes fueron acogidos. Pero lo que sus arquitectos no entendieron era que todas las comunidades —sobre todo aquellas que carecían en buena medida de reglas o autoridades formales— se autogobiernan por costumbres sociales de algún tipo. Y las normas y valores que ellos habían codificado en los inicios de la web resultaron guiar a los millones de personas que la adoptaron en los primeros años hacia algo muy alejado de la utopía igualitaria que ellos habían imaginado.

En los albores de lo que ahora llamamos cultura de internet, allá por 2009, un chaval de trece años que vivía en un barrio residencial de Dallas llamado Adam inició un viaje por las profundidades de la web social que le ocuparía la mayor parte de su vida. Había oído en alguna parte la palabra *meme*, la había buscado en Google y había ido a parar a un blog llamado *I Can Haz Cheezburger?* («¿Me pondrá una hamburguesa con queso?»). La página atraía a dos millones de usuarios diarios, que veían fotos de gatos con textos infantiles superpuestos que al parecer se leían con voz de gato («i maked you a dinner but i eated it»), «te preparé la cena pero me la comí»).

A Adam le llamó la atención una publicación —una serie de memes y capturas de pantalla— que contaban una historia. La publicación decía que habían aparecido en YouTube un par de vídeos en los que se veía a dos niños abusando de un gato llamado Dusty.³⁰ Los vídeos generaron indignación, sobre todo en un tablón de

mensajes llamado 4chan, cuyos usuarios podían ser malévolos. Prometieron justicia por Dusty; uno identificó al creador del vídeo de YouTube como un chico de catorce años de Oklahoma. Otro afirmó que había avisado a la policía, si bien más adelante las autoridades dijeron que habían identificado el origen del vídeo por su cuenta. El creador del vídeo y su hermano fueron acusados de actuar con crueldad con los animales; 4chan se llenó de publicaciones en las que se celebraba la victoria como un éxito propio.

Adam estaba embelesado. Personas como él habían convertido el acto de hacer clics por internet en una aventura emocionante y una experiencia social de vinculación afectiva. «¿Que unos tíos salvaron a un gatito? Pues claro que aquello me atrajo —recordaba el chico—. El espíritu justiciero contribuyó a definir lo que 4chan es para mí.» Entonces empezó a pasar varias horas diarias en 4chan, cuyas inusuales características lo habían hecho de lo más popular entre los primeros usuarios de internet.³¹ Cuando un usuario quería abrir un nuevo hilo, tenía que subir una imagen, lo cual hacía que la plataforma siempre estuviese llena de memes y dibujos hechos por los usuarios. Mucho antes de que Snapchat y otras plataformas copiaran esa función, los debates se borraban de forma automática al cabo de poco tiempo, lo cual permitía un comportamiento indecoroso que en otros sitios probablemente se habría evitado. Y a eso mismo contribuía el anonimato del portal; casi todas las publicaciones aparecían con la etiqueta que indicaba que estaban escritas por «Anónimo», lo cual infunde una cultura del todo vale y un sentimiento de identidad colectiva que puede ser atractivo, sobre todo para personas que ansían sentirse parte de algo.

«En el colegio, yo siempre fui el marginado», dijo Adam durante una de nuestras conversaciones a altas horas de la noche sobre los rincones más oscuros de internet, por los cuales se había ofrecido a guiarme. Él había sufrido depresión y ansiedad desde niño. Pero en 4chan se sentía acogido y comprendido. Noche tras noche, sus usuarios se entretenían unos a otros con hilos escandalosos y chistes rebuscados que eran «de las cosas más divertidas que he visto por internet», aseguró Adam.

Y allí se sentía seguro de una forma en la que no se sentía en el mundo real, donde, según admite él mismo, era «una persona bastante rara» y «un poco encerrada en sí misma». El dormitorio que tiene en

casa de su madre, donde aún vive en mitad de la veintena, estaba «cubierto», en palabras suyas, de «artículos promocionales» de un videojuego de los noventa. Mientras que mencionar su depresión en el mundo real generaba miradas de preocupación y la intervención de los adultos, en 4chan podía abrirse a otras personas que parecían compartir su soledad.

En una foto suya que me mandó, llevaba una camiseta retro de un videojuego y unas gafas negras a la moda medio escondidas debajo de unas greñas castañas. No parecía alguien que necesitase refugiarse en unas complejas redes sociales para hacer amigos o conocer chicas. Pero es tímido en persona, me dijo, porque cecea y tartamudea. Ambos rasgos se vieron agravados por una pérdida de audición cuando era adolescente y, más recientemente, por ansiedad. Aún le cuesta la comunicación verbal y no trabaja. «Soy un NEET, tío», dijo usando un acrónimo del Gobierno británico para alguien que «ni trabaja, ni estudia ni se está formando»,³² que en el mundo virtual se ha convertido en una etiqueta con la que se identifican aquellas personas que se sienten abandonadas.

Al igual que muchas personas que han residido durante largo tiempo en la web social profunda, él era reticente a hablar con periodistas, a quienes veía como representantes del *establishment* adulto que quería, según él, controlarlo y clausurar su amado hogar digital. Pero un intermediario de confianza, una fuente mía con quien él había contactado unos diez años antes después de unirse a 4chan, nos puso en contacto. A pesar de que Adam había disfrutado de los foros y de su «diversión en su mayor parte anónima y banal», admitió que le preocupaban «los rincones oscuros donde el comportamiento antisocial se filtra al mundo real», y añadió: «Lo vi desde el principio. Soy judío y la mitad del portal cree que no debería existir». Le costó entender cómo aceptar por completo la visión del mundo que saturaba la plataforma. Su cultura de la rabia y la conspiración le ayudó a encontrarle sentido a un mundo confuso que parecía que iba en su contra. El hecho de que los agravios fuesen colectivos le ayudó a tener un sentimiento de pertenencia. Pero los extremos descontrolados le preocupaban, sobre todo cuando se hacía daño a alguna persona.

«En última instancia —dijo Christopher Poole, creador de 4chan, en 2008—, la comunidad tiene el poder de dictar sus propios estándares.»³³ Poole, que entonces era un chaval flacucho de veinte

años, pertenecía a una creciente comunidad de aficionados a internet —programadores, blogueros, estudiantes de posgrado— que estaban construyendo la «civilización de la Mente» de Silicon Valley. Pero, después de todos los emocionantes congresos de *hackers* de los años ochenta y los manifiestos de los noventa, en la década del año 2000 los habitantes de la web estaban interesados antes que nada en pasarlo bien. Adolescentes de barrios residenciales, sin poder salir de casa después del colegio, demasiado jóvenes para conducir, se pasaban horas en internet. La promesa de libertad absoluta que les brindaba la red era especialmente atractiva para esos chavales, para quienes la vida no virtual está determinada por los padres y los profesores. Los adolescentes también tienen un impulso más fuerte que los adultos para socializar,³⁴ lo cual se manifiesta en un uso mayor de las redes sociales y una mayor sensibilidad ante lo que allí sucede.³⁵ Poole había creado 4chan cuando tenía tan solo quince años. Los chicos que se sentían aislados en el mundo físico, como Adam, representaban un porcentaje desproporcionado de la actividad de internet y trajeron consigo las preocupaciones de los marginados y los acosados.

Esa cultura supuso, de entrada, creatividad y banalidad —vídeos y dibujitos de gatos—, pero también tendía a ser traviesa y transgresora. Retomando la cultura de la web que se remontaba al WELL, la diversión significaba estar libre de las normas y las sensibilidades de la sociedad y despreciar a los ajenos a ese mundo que no pillaban los chistes. Así fue como el hecho de gastar bromas —un pasatiempo que gusta a los hombres jóvenes de todas partes— se convirtió en la actividad definitoria de los inicios de la web.

En una artimaña organizada en 4chan, los usuarios secuestraron un concurso virtual en el que los colegios competían para ganar la organización de un concierto de Taylor Swift y dirigieron la victoria a un centro para niños sordos. En otro caso, los usuarios colgaron una broma interna de 4chan, salida de alguna sesión de memes a altas horas de la noche, en un tablón de mensajes para el programa televisivo de Oprah Winfrey, donde escribieron: «Nuestro grupo tiene más de nueve mil penes y son todos para violar niños». Winfrey citó la frase con aire serio en directo. «Gran parte del humor —dijo Adam— proviene del acto de ensanchar los límites y extender su alcance hasta sitios más mayoritarios que parecen intocables.» Meterle un gol a un mundo que los trataba como a unos apestados.

Las situaciones a veces derivaban en puro sadismo. En 2006, los usuarios descubrieron que, tras el suicidio de un chico de trece años, los amigos de la víctima habían publicado comentarios de recuerdo en su página de Myspace. Los usuarios se burlaron de los errores ortográficos y de la sinceridad infantil. Algunos hackearon la página y cambiaron la foto de perfil del chico por la imagen de un zombi. Otros hicieron llamadas de broma a sus padres, lo cual se prolongó durante más de un año. Si a los adultos y a los compañeros de clase esas cosas les parecían desagradables, pues mucho mejor.

Transgredir tabúes cada vez mayores —incluso actuar con crueldad contra unos padres que han perdido a un hijo— se convirtió en una forma de señalar que estabas en el ajo. «Cuando navegas por 4chan y 8chan mientras el resto de tus amigos están colgando las mierdas típicas de fiestas y parejitas en Instagram y Facebook —dijo Adam— te sientes distinto. Más guay. Como que eres parte de algo friqui.» La broma podía ser una foto de pornografía escatológica. Un vídeo de un asesinato macabro. Un insulto racista publicado para fastidiar a la gente y retarla a tomárselo en serio. Reírse del material —o, mejor incluso, colgar algo aún más fuerte— afirmaba que compartías la indiferencia cínica y cómplice del club. Y eso reconfiguraba la relación de sus miembros con el mundo exterior: no es la sociedad quien nos rechaza, sino nosotros a ella.

Estas dos actividades unificadoras —romper tabúes y gastar bromas— convergieron y se convirtieron en el troleo. Un tablón de mensajes de los años noventa había definido el troleo como la publicación de comentarios «con el mero objetivo de importunar a alguien o de entorpecer un debate»,³⁶ llamado así probablemente por la palabra inglesa *troll*, que hace referencia a «un estilo de pesca en el que el pescador arrastra el cebo por un caladero plausible con la esperanza de que piquen». Desde la época del WELL, los usuarios de internet se habían entretenido intentando provocarse unos a otros. En redes como 4chan, eso a menudo se convirtió en algo más turbio: actos de abuso colectivo en los que el objetivo era deleitarse con la angustia de otra persona.

La emoción de sacarle una reacción a otra persona incluso tenía un nombre: *lulz*, una corrupción del acrónimo LOL («laugh out loud», ‘reírse a carcajadas’). Casi no había ningún límite en cuanto a quién podía ser objeto de burlas y al grado de agresividad de las acciones; el

sueño de Silicon Valley de liberarse de leyes y jerarquías se había vuelto, en internet, una libertad también con respecto a los códigos sociales y morales. La comunidad había creado sus propios estándares, como pretendía Poole, en torno al incentivo fundamental de todas las redes sociales: la atención.

La mayor parte de las plataformas están construidas alrededor de una creencia aparentemente neutra de que atención indica valor. Reddit y Twitter colocan en una posición más destacada las publicaciones basándose en cuántos usuarios les dan votos positivos o retuits. Facebook y YouTube incorporan esta autoridad en los algoritmos. Cada versión unifica todas las formas de atención — positiva o negativa, irónica o sincera, que haga reír o que haga enfadar, que satisfaga tu curiosidad intelectual o tus instintos lascivos — en un solo indicador: arriba o abajo.

Los tabloncillos de mensajes como 4chan hacen lo mismo, aunque de un modo más orgánico. El anonimato y la producción de contenido animan a los usuarios a buscar la atención de los demás con la mayor agresividad posible. Sin las ataduras sociales del mundo físico, cada usuario opera como un algoritmo de Facebook en miniatura y aprende por repetición lo que mejor sirve para captar la atención de los demás. Es una lección que funciona de un modo sistemático. Para destacarse en medio de decenas de miles de voces, con independencia de cómo sea la publicación, es mejor subir el volumen, ser más extremo.

En 2010, uno de los ataques más aplaudidos de la comunidad estuvo dirigido a una niña de once años.³⁷ Los usuarios habían visto publicaciones que circulaban por Myspace, donde la niña tenía un perfil, en las que se afirmaba que ella había dicho a algunos amigos que mantenía una relación sentimental con un chico de veinticinco años que era el líder de un grupo de música local. Los usuarios de Myspace se burlaron de ella por —según suponían ellos— haberse inventado la historia. Ella colgó respuestas airadas, luego un vídeo de YouTube diciendo a sus *haters* o «detractores» que les «reventaría la cabeza de un pistoletazo».

Los usuarios de 4chan, deleitándose con la reacción torpe y emotiva de la niña, siguieron con sus provocaciones. La niña había mentido, no había asumido responsabilidades y ahora la pondrían en su sitio. La bombardearon con mensajes de acoso en Facebook y Myspace, falsas entregas de pizza a domicilio y llamadas de broma, las

cuales grabaron y colgaron en 4chan para divertirse más aún. Una publicación sobre «Cómo troleear» a la niña aconsejaba: «Decidle que se suicide» y «Decidle a su padre que vamos a pegarle una paliza a su hija» y a distribuir fotos íntimas suyas.

La niña publicó un segundo vídeo, en el que sollozaba mientras su padre adoptivo gritaba a sus invisibles torturadores: «Ya veréis: pillarán vuestros emails y os van a encontrar». Aquello entusiasmó a los usuarios de 4chan; era una prueba de que estaban consiguiendo su objetivo. El acoso empeoró. En un vídeo final, la niña cedía, admitía todas las acusaciones y pedía misericordia. El incidente se volvió legendario en la red social: justicia y *lulz*, conseguido. Los vídeos de la niña se difundieron más y más, lo cual atrajo millones de visualizaciones. El grupo de música al que había nombrado en sus publicaciones iniciales grabó una canción que aplaudía su humillación: «Mi nombre y mi reputación no serán el objetivo de una guerra / Yo estaré tocando el cielo y tú estarás jodida cortándote las venas».³⁸

Años después, la niña contó a la policía que, cuando tenía diez años, el líder del grupo la había acosado sexualmente y violado varias veces. Como ocurre con muchas víctimas de abusos infantiles, la vergüenza y la confusión la habían llevado a contarles a sus amigos que las relaciones eran consentidas. En el marco de una investigación del FBI, salieron entre veinte y treinta personas más con historias parecidas.³⁹ Aun así, el episodio sigue festejándose en 4chan y en sitios similares, donde los usuarios han continuado acosándola en toda la red durante más de una década.

La desvergüenza era un rasgo característico; los usuarios de 4chan «solían plantear su actividad como una suerte de servicio público»,⁴⁰ escribieron en una ocasión los teóricos de la cultura Whitney Phillips y Ryan Milner. Creían que participaban en la misión fundacional de la web social: acabar con los comportamientos anticuados y sustituirlos por un mundo de libertad de expresión autorregulado y con un pensamiento despiadadamente independiente. Engañar a Oprah para que alertase a sus espectadores de una secta sexual inventada, se contaban a sí mismos, enseñaría a los ciudadanos a cuestionar a los personajes de autoridad y los relatos oficiales, «un resultado que se esforzaban por conseguir mediante burlas, manipulaciones y engaños —escribieron Phillips y Milner—. Muchos

bromeaban con que merecían un agradecimiento por sus acciones».⁴¹

«El troleo es, básicamente, eugenesia virtual»,⁴² dijo en 2008 Andrew Auernheimer, una superestrella de 4chan: *hacker* anarquista, provocador descarado y torturador despiadado de los enemigos elegidos del portal. La entrevista en sí ya era un troleo, y se provocaba a los lectores para que se ofendiesen con invocaciones nazis: «Los blogueros son escoria [...]. ¡Debemos meter a esa gente en el horno!». Las provocaciones como las de Auernheimer apuntaban a algo nuevo que estaba surgiendo desde los inicios de la web social. La cultura de la provocación, el alegre rechazo de los códigos morales establecidos, había creado lo que Phillips y Milner llamaron «las condiciones perfectas para que la intolerancia se propagase con sigilo, escondida dentro de cosas que no parecían nada perversas».⁴³

Auernheimer se jactó de su papel en una campaña de acoso contra Kathy Sierra, una bloguera sobre tecnología.⁴⁴ Sierra había pedido que se moderaran las secciones de comentarios de 4chan, lo cual indignó a los usuarios, que vieron aquello como un ataque a la libertad de expresión en internet. Auernheimer publicó el número de la Seguridad Social de Sierra y su dirección; y la mujer recibió una avalancha de amenazas de muerte, se divulgaron fotos de sus hijos editadas como imágenes pornográficas y se publicaron mensajes en su nombre ofreciendo sexo, con lo cual se presentaban a su casa hombres extraños a todas horas. Sierra se vio obligada a desaparecer por completo de la vida pública. La mujer tenía pocos incentivos para intentar presentar acusaciones. En el sistema judicial, pocas personas consideraban que el acoso por internet, aunque fuera extremo, cruzase el umbral de la criminalidad, y, sin duda, cualquier respuesta pública no habría hecho sino provocar más ataques.

Auernheimer también empezó a colgar «sermones» conspiratorios y antisemitas en YouTube que afirmaba que eran satíricos, pero cuyas llamadas a que hubiese «sangre en las calles» sonaban terriblemente sinceras. En ese momento, a los Auernheimers de la red se les trataba como unos chavalines traviesos. Y si alguien les prestaba atención, a menudo era para elogiarlos. En 2010, *TechCrunch*, una publicación interna casi oficial de Silicon Valley, concedió a Auernheimer un premio por su servicio público.⁴⁵ Era una muestra de solidaridad con el grupo de *hackers* de Auernheimer, que estaba siendo investigado por el FBI por robar 114.000 direcciones de correo privadas a través de

una brecha de seguridad de un iPad.

En una fiesta en honor de Auernheimer celebrada la noche antes de anunciarse su condena penal, «los periodistas bebieron junto a *hackers*, activistas, documentalistas con un parche en el ojo y chicas con el pelo de colores chillones y *piercings* en los labios»,⁴⁶ según informó *Vice*. A la mañana siguiente, el redactor de *Vice* trajo una máscara de Guy Fawkes al juzgado y explicó: «Vine porque su condena es un error, y mis amigos y yo nos preocupamos por él». Al año siguiente, Auernheimer empezó a trabajar para el *Daily Stormer*, un destacado foro neonazi fundado y habitado por usuarios de 4chan, donde él colgó una foto en la que mostraba un tatuaje de una esvástica del tamaño de un puño que llevaba en el pecho.

Poole, preocupado por la reputación cada vez más turbia de 4chan, impuso la restricción más liviana posible. Seguían estando permitidos los discursos de odio extremo y el acoso, pero quedaban limitados a algunas subsecciones. El creador del tablón convirtió esas secciones en barrios rojos del troleo, lo cual atrajo un aluvión de mirones, a algunos de los cuales les gustó lo que vieron y se quedaron. A pesar de todo, hubo quienes consideraron que eso también era una traición. Una de esas personas, un programador llamado Fredrick Brennan, creó una nueva versión de 4chan (8chan), que anunció como una alternativa «que sí respetaba la libertad de expresión». Los usuarios, entre ellos Adam, se trasladaron en tromba a la nueva plataforma, lo cual reforzó la identidad colectiva de unos *outsiders* desafiantes que acogería algunos de los mayores horrores de la era de internet.

Al terminar la década de los 2000, las comunidades chan migraron a las plataformas estéticamente más atractivas que estaban en auge entonces. Como si fueran oleadas de colonos, aterrizaron en las tierras vírgenes de Facebook, YouTube y Twitter. La cultura chan, expresada en memes y chistes internos ahora tan reconocibles como Mickey Mouse, llenaron las plataformas, cuyas características y algoritmos para la maximización de la participación absorbieron y cosificaron sus tendencias más extremas, y luego las amplificaron a un mundo que no tenía ni idea de lo que se le estaba viniendo encima.

Cuando John Doerr, el influyente inversor tecnológico, inauguró a finales de 2010 el fondo Kleiner Perkins, dotado con 250 millones de dólares para empresas emergentes de redes sociales, la persona a la que eligió para estar a su lado fue Bing Gordon. Gordon, que se describía a sí mismo como un jugador de *lacrosse*, entonces en la cincuentena, llevaba unas greñas de universitario y adornaba su argot corporativo con palabras como «tío» y «superguay». Antes de unirse a Kleiner como inversor, su *curriculum vitae* tenía un apartado que abarcaba de 1982 a 2008: Electronic Arts, una empresa de videojuegos. «Estamos apostando por un sector que está por explorar, y creemos que las redes sociales no han hecho más que empezar»,⁴⁷ dijo Gordon el día de la inauguración, flanqueado, en una muestra del poder de Kleiner, por Mark Zuckerberg y Jeff Bezos.

La presencia de Gordon reflejaba una creencia muy generalizada entonces: el sector de las redes sociales operaría como el negocio de los videojuegos. Unos meses después, Gordon dijo en un congreso del sector que había «tres temas que vosotros, los directores ejecutivos, debéis dominar»,⁴⁸ y enumeró los móviles, las redes sociales y, haciendo un gesto de victoria con el puño, la «gamificación». Desde el principio, las plataformas de redes sociales tomaron muchos elementos de los videojuegos. Las notificaciones se muestran en forma de «recuadros» estilizados, lo cual —según Gordon le contó al público— podía doblar el tiempo que un usuario se pasaba en el sitio, mientras que los «me gusta» imitan un marcador que va subiendo. Aquello iba más allá de la estética. De entrada, muchas plataformas consideraron a los *gamers* —personas obsesionadas con la tecnología que seguro que también dedicarían muchas horas a esa interfaz digital— como un mercado fundamental.

Gracias a una casualidad de la historia del comercio, el sector de los videojuegos se dirigía en un porcentaje muy mayoritario a hombres jóvenes y chicos de ciertos temperamentos, lo cual conllevó que las plataformas de redes sociales hicieran lo mismo. Pero en los videojuegos no hay nada que sea inherentemente de un género o una edad concreta. Los primeros juegos, diseñados por empresas de Silicon Valley de los años setenta como Atari, se pusieron a la venta junto a los ordenadores personales y se suponía que también se dirigían a todos los públicos. Eso cambió con lo que el sector denomina el crac de los videojuegos en América del Norte.⁴⁹ Entre 1983 y 1985, las

ventas disminuyeron un 97 %. Las empresas japonesas intentaron reavivar el mercado renovando la imagen de aquel producto informático, ahora caído en desgracia, que se vendía a adultos en tiendas de electrónica, como algo más simple: como juguetes.

En ese momento, las secciones de juguetes estaban muy segmentadas por género. El presidente Reagan había levantado las restricciones que prohibían la publicidad en televisión destinada a niños. Los equipos de *marketing*, impulsados por un neofreudianismo que entonces estaba en boga, creyeron que podían atraer a los niños satisfaciendo su curiosidad naciente por su propio género. Nuevos programas televisivos como *My Little Pony* y *GI Joe* transmitían unas normas de género hiperexageradas, lo cual influía en el autodescubrimiento natural del género por parte de los adolescentes y lo convertía en un deseo de tener productos de plástico moldeado. No es casualidad que eso suene como un eco sorprendentemente nítido del modelo de negocio de las redes sociales. Sacar provecho de nuestras necesidades psicológicas más profundas, luego enseñarnos a tratar de saciarlas mediante un consumo comercial que nos dejará insatisfechos y deseosos de más, ha sido un elemento fundamental del capitalismo estadounidense desde el *boom* de posguerra.

Las secciones de juguetes polarizaron entre rosa y azul. Los fabricantes de juegos japoneses tuvieron que elegir un bando, así que seleccionaron aquel en el que los padres gastaban más dinero: el de los chicos. Los juegos se centraron cada vez más en héroes masculinos que rescataban princesas, que combatían en guerras y que jugaban a ligas de deportes masculinas. Los responsables de *marketing*, que ya llevaban tiempo posicionando los juegos como juguetes para niños, mantuvieron a los chavales enganchados durante la adolescencia y la edad adulta con —¿qué iba a ser?— el sexo. Llenaron los juegos de personajes femeninos representados como seres hipersexualizados, sumisos y algo a lo que los hombres deberían sentir que tienen derecho. Muchos *gamers* entendieron que aquellas representaciones eran fantasía, aunque fuese una fantasía con unos valores problemáticos. Pero buena parte de ellos crecieron con esa fantasía y la absorbieron como una verdad. En medio de las guerras culturales de las décadas de 1990 y del año 2000, los responsables de *marketing* de las empresas de videojuegos aprovecharon esas tendencias como un activo y presentaron los juegos como un refugio ante un mundo que se

estaba feminizando, un lugar donde los hombres seguían siendo hombres y las mujeres se quedaban en su sitio. Los videojuegos se convirtieron, para algunos, en una identidad, una identidad anclada en una reacción contra la evolución de las normas de género.

Cuando aparecieron las redes sociales, accesibles solo desde los ordenadores de sobremesa colocados en el salón de las casas, los primeros usuarios se parecían mucho a los *gamers*: hombres y chicos caseros que fueron de los primeros en adoptar las novedades tecnológicas. De ahí que los dos públicos fueran tratados como sinónimos, y los productos que eran para unos se intentaban vender también a los otros, lo cual fue juntándolos en una subcultura e identidad unificadas. Pero, al comenzar la década del año 2000, unos ordenadores cada vez más potentes permitieron a programadores autofinanzados crear juegos fuera del molde comercial. (Zoë Quinn fue uno de ellos.) La democratización digital también trajo nuevas voces, como la *youtuber* Anita Sarkeesian, quien argumentó que las representaciones de las mujeres en los videojuegos no solo las excluían, sino que fomentaban que se las maltratase en la vida real. Entre los *gamers* deseosos de que se tomasen más en serio los patrones representados, aquella evolución fue emocionante y positiva. Los críticos empezaron a emplear los argumentos de Sarkeesian. Los estrenos de gran presupuesto presentaban protagonistas y tramas diversos, a la vez que reducían los excesos de masculinidad.

No obstante, algunos hombres jóvenes *gamers* —sobre todo aquellos que se reunían en las plataformas sociales— vieron aquello como una amenaza. Las normas y los límites de su identidad estaban en peligro y, con ello, la percepción que tenían de sí mismos. «Durante treinta años tuvimos un tipo muy concreto de jugador; y lo hemos convertido en el centro del mundo y hemos satisfecho todos sus antojos —me dijo Brianna Wu, que había osado cuestionar a los *gamergaters*—. Y ahora, hoy, el mundo está cambiando, y el jugador medio no es un chico de veintitantos años. Es una mujer de cuarenta como yo. Y ante este cambio, simplemente vemos una incapacidad para aceptarlo que es muy desafortunada en nuestro ámbito.»

Acusar a programadores como Quinn de sobornar a críticos era una suerte de autodefensa cognitiva, una forma de presentar un cambio indeseado como una conspiración infame, y la amenaza a su identidad como una batalla contra ellos. «Ya no estaban luchando por

el derecho a ver tetas en los videojuegos, sino que luchaban contra el “genocidio blanco” —ha dicho David Futrelle, un autor que monitoriza extremistas virtuales—. No solo había que salvar los videojuegos, sino la propia civilización occidental.»⁵⁰ Echando la vista atrás, Futrelle vinculaba ese hecho al incremento explosivo de la extrema derecha en internet, y añadía que era «el Gamergate había tenido, sin lugar a dudas, una gran importancia en el proceso de radicalización».

Adam, el obseso de 4chan, que ahora tenía dieciocho años, siguió todas las huellas. Aquella indignación en que se afirmaba su visión del mundo era irresistible. Los vídeos colgados por destacados *youtubers* sobre videojuegos lo convencieron, dijo, de que «a los periodistas los estaban sobornando con dinero y sexo». No le gustaba ver a algunos usuarios planear cómo atormentar a Wu. Pero las plataformas, aprovechando su ingeniería, lo bombardearon con publicaciones que supuestamente demostraban que Wu y otras personas eran radicales peligrosos empeñados en someterlo. Continuó clicando, siguiendo la campaña, como muchos de sus compatriotas virtuales, y fue cayéndose por madrigueras aún más oscuras y profundas.

A Wu, al ver la evolución del Gamergate, le vino a la cabeza un episodio de años antes, cuando había sido becaria en la oficina del senador por Misisipi Trent Lott. Su equipo había realizado una encuesta interesada que con el tiempo se ha hecho famosa: «¿Cree usted que los demócratas están intentando arrebatarle su cultura?». La encuesta dio unos resultados espectaculares, sobre todo entre los hombres blancos. Imagínate, dijo Wu, lo efectivas que podrían ser las redes sociales si se optimizasen para desencadenar las emociones de las personas incluso con mayor eficacia que el equipo de campaña más astuto; podrían saturar a miles de millones de personas con una versión de la realidad que fuese como una permanente encuesta interesada que activase su identidad. «Justo de la misma forma que esa encuesta interesada de Misisipi, creo que había un miedo real a que las mujeres estuviesen arrebatándoles su cultura —dijo sobre el Gamergate—. Es algo tribal.»

Durante los primeros años de la década de 2010, cuando las subculturas del estilo del Gamergate infestaron las plataformas de primera generación, Facebook perseguía con discreción un objetivo osado: superar los límites cognitivos de la socialización humana. El plan, con el tiempo, impondría, a miles de millones de usuarios, una versión aún más potente de las distorsiones alteradoras de la mente que Reddit y 4chan habían traído a los *gamers* y a los primeros que tuvieron perfil en sus redes, lo cual desembocó en un mundo de Gamergates.

Aquello surgió a raíz de una crisis que la empresa sufrió en 2008, una crisis que captó la atención de la empresa como pocas: se había estancado el crecimiento del número de usuarios. En cualquier otro sector, alcanzar cerca de noventa millones de clientes podría haber sido una oportunidad para explorar productos nuevos o mejores que venderles. Pero, en la economía de internet, una base de usuarios estática podía ser mortal. «Recuerdo que algunas personas decían que en ese momento no estaba claro que algún día pudiésemos superar los cien millones de usuarios —ha dicho Zuckerberg—. En definitiva, topamos contra un muro y teníamos que abordar aquella cuestión.»⁵¹

Facebook, con la esperanza de fomentar la participación, comenzó a experimentar con la ruptura del llamado límite de Dunbar. En los años noventa, el antropólogo británico Robin Dunbar propuso que los humanos estaban limitados en términos cognitivos al mantenimiento de unas 150 relaciones.⁵² Ese número derivaba de los grupos sociales, de un máximo de 150 personas, en los que habíamos evolucionado. Si traspasamos ese límite, nuestro neocórtex —la parte del cerebro que controla la cognición social— se agota. Incluso en el mundo virtual, la gente convergía de forma natural en el número de Dunbar. En 2010, el usuario medio de Facebook tenía unos 130 amigos;⁵³ el videojuego social Friendster incluso limitó el número de amigos a 150.⁵⁴

Durante mucho tiempo, Silicon Valley había soñado con «escapar de la maldición de Dunbar»,⁵⁵ como escribió en 2010 un consultor de las empresas del valle. Zuckerberg habló en público de acabar con la maldición.⁵⁶ Pero el plan no salió bien. Ni siquiera Facebook podía superar millones de años de evolución; al menos, no todavía. Cuando el crecimiento del número de usuarios volvió a estancarse en 2013, la compañía se reestructuró en torno a ese objetivo. Dirigió a los

usuarios hacia contenido de lo que Facebook llamaba «vínculos débiles»: amigos de amigos, contactos de contactos, primos de primos.

Mediante sofisticados algoritmos, la estrategia funcionó. Facebook empujó a los usuarios hacia círculos cada vez mayores formados por individuos medio desconocidos y así superó el límite de Dunbar. Twitter, en esa misma época, hizo más o menos lo mismo: empezó a mostrar a los usuarios tuits de desconocidos y a animarlos a seguir a amigos de amigos. Las compañías mostraron poca preocupación por las consecuencias de saltarse nuestros límites neurológicos. Operaron, como siempre habían hecho, creyendo que sus productos eran liberadores por naturaleza.

No obstante, en algunos estudios sobre los monos y macacos Rhesus,⁵⁷ cuyos límites de relaciones sociales se cree que se parecen a los nuestros, se había llegado a la conclusión de que forzar a los individuos a relacionarse en grupos más grandes hacía que se volvieran más agresivos, más desconfiados y más violentos. Era como si se amplificasen todos los peligros de vivir en una comunidad y se redujeran los placeres. Los monos parecían darse cuenta de que navegar con seguridad entre un grupo de un tamaño antinaturalmente grande excedía sus capacidades, lo cual desencadenaba una respuesta de lucha o huida social que nunca se desactivaba por completo. También parecían estar más confusos al formar e imponer jerarquías sociales, probablemente como una suerte de mecanismo de defensa.

Facebook pronto descubrió un método aún más potente para ampliar sus comunidades de usuarios. En lugar de esforzarse por ampliar tu lista de amigos más allá de las 150 personas, Facebook podía empujarte a entrar en grupos —páginas de debate independientes centradas en algún tema o interés— de un tamaño diez veces superior. Eso también dio aún más poder a los sistemas de Facebook. Como ya no estaban limitados al contenido de personas cercanas a tu círculo social, el sistema podía recomendarte grupos de cualquier rincón de la plataforma.

Renée DiResta, la inversora tecnológica que había descubierto el crecimiento de las comunidades antivacunas en Facebook, se dio cuenta de la importancia de ese cambio cuando su batalla entró en una nueva fase. En California se debatía sobre una ley que endurecería la obligación de vacunarse en los colegios. En internet, ella apoyó el proyecto de ley y la ciencia que había detrás con la esperanza de

contrarrestar la opinión antivacunas tan predominante en el mundo virtual. Facebook empezó a recomendarle que se uniera a otros grupos de padres. Luego a grupos antivacunas, aunque eso no parecía sorprendente, teniendo en cuenta su actividad. Pero al poco tiempo Facebook le ofreció unirse a grupos basados en la desinformación médica sobre otros asuntos. Muchos grupos difundían una conspiración de la que ella nunca había oído hablar: que el zika, un virus que se estaba propagando por América Latina y Estados Unidos, había sido creado de forma artificial. Buscó en Facebook el término *zika* para comprobar si, como ocurría con las vacunas, la plataforma podía estar empujando a los usuarios hacia los extremos. Como era de esperar, los mejores resultados estaban todos relacionados con grupos conspiratorios que decían que el zika era un complot judío, un plan para controlar a la población o el primer paso para que alguien se hiciera con el poder mundial.

«Hay un efecto de correlación entre conspiraciones —dijo DiResta— en el que la plataforma reconoce que alguien que está interesado en la conspiración A es probable que esté interesado también en la conspiración B; y, por tanto, le aparece como recomendación.» La época de los grupos de Facebook fomentó algo más específico que el consumo pasivo de conspiraciones. Leer un texto sobre las estelas de los aviones o sobre los virus creados en un laboratorio puede ocupar veinte minutos. Pero unirse a una comunidad organizada en torno a combatir algo podía convertirse en una rutina diaria durante meses o años. Cada vez que un usuario sucumbía, se entrenaba al sistema para que impulsase a otros a hacer lo mismo. «Si muerden el anzuelo —dijo DiResta—, entonces han reforzado ese aprendizaje. Luego el algoritmo tomará ese refuerzo y le dará más peso.»

Otros miembros del grupo informal de DiResta dedicado a observar las redes sociales se estaban dando cuenta de que Facebook y otras plataformas los estaban encaminando de formas parecidas. Se repetía una y otra vez el mismo patrón, como si aquellas inteligencias artificiales hubiesen llegado todas por separado a una verdad común y terrible sobre la naturaleza humana. «Yo lo llamé radicalización mediante el motor de recomendaciones —dijo DiResta—. A partir de unos indicadores derivados de la participación en la plataforma, creas un mundo en el que el contenido que es fruto de la indignación se convierte en la norma.»

La lógica algorítmica era sensata, brillante incluso. La radicalización es un proceso obsesivo y que consume mucho tiempo. Los creyentes vuelven una y otra vez, su obsesión se va convirtiendo en una identidad y las redes sociales en el centro de su vida cotidiana. Y los radicales, impulsados por la urgencia de su causa, reclutan más radicales. «Habíamos construido una máquina de la ira en la que las personas participaban en la difusión del contenido», dijo DiResta, donde a partir de entonces los que se radicalizaban eran «los diseminadores de ese contenido». Ella lo había visto miles de veces. A los reclutados los unía una presunta amenaza de vida o muerte: la terrible verdad de las vacunas, los agentes de los *illuminati* que propagaban el zika, las feministas que pretendían acabar con el lugar legítimo de los hombres en el escalafón más alto de la jerarquía de género, empezando por los videojuegos. «Personas normales y corrientes empezaban a sentirse soldados de un ejército virtual que luchaba por su causa», me explicó DiResta. Era solo cuestión de tiempo que desearan pasar a la acción entre todos.

Sin embargo, mientras que DiResta podía demostrar que eso ocurría en casos concretos, como no tenía acceso a la inmensa cantidad de datos internos de la empresa, solo podía inferir que las plataformas estaban fomentando eso de forma sistemática. Avisó a sus compañeros de Silicon Valley —primero dentro de su grupo informal de observadores, luego más en general, en artículos y conferencias públicas— de que algo no estaba funcionando bien. Le preocupaba, no obstante, que fuera demasiado tarde. «Tenía la sensación —me dijo— de que daba igual que fuesen aficionados a los videojuegos, antivacunas» o cualquier otro grupo que se estuviese radicalizando a gran escala. «Esa era la dinámica que se estaba formando como consecuencia del sistema.»

Desde este punto de vista, el Gamergate parecía casi inevitable; y eso es justo lo que sucedería cuando los primeros que se habían metido en las redes sociales —los *gamers*— interactuasen con una máquina calibrada con pericia para producir una respuesta como esa. Comunidades más grandes, más convencionales, de personas normales y corrientes sin una especial vinculación con internet, ya estaban adentrándose en el mundo virtual y entregándose a la influencia de la máquina. Había muchos motivos para suponer que su experiencia sería muy similar. En los primeros meses de 2015, eso ya estaba

empezando.

Abriendo el portal

LA LLAMADA DE ALERTA

«Creo que no sabían muy qué pensar de mí —recordaba Ellen Pao sobre sus primeros días en Reddit—. La bienvenida no fue superacogedora.» En la primavera de 2013, Pao era una veterana de Silicon Valley que había trabajado durante un tiempo en la prestigiosa empresa de inversión Kleiner Perkins. Pero, a pesar de su paso por Kleiner, que cuenta con un enorme fondo para empresas emergentes de redes sociales, Pao tenía poca experiencia directa con las compañías de ese ámbito. Y destacaba en otro sentido: era una mujer no blanca de alto rango empresarial en un sector tristemente conocido por estar dominado de forma abrumadora por hombres.

Pao, conocedora de Silicon Valley tanto desde fuera como desde dentro, creía firmemente en ese lugar desde los años noventa, cuando había empezado a trabajar para una empresa que fabricaba un *hardware* que permitía acceder a internet desde el televisor. Aun así, cada vez era más escéptica con respecto a la inclinación de las grandes tecnológicas hacia los hombres jóvenes *geeks*. Aparte de Yishan Wong, el director ejecutivo de la empresa, los veintipico empleados de la flamante Reddit eran en su mayoría jóvenes. Buena parte de ellos hablaban usando las bromas internas y el argot cultural de una plataforma que, a pesar de tener una cantidad enorme de usuarios, daba la sensación de ser un sitio aislado. Pao era un poco fría; Wong la instó a usar un lenguaje menos «empresarial».

Además, no era una gran usuaria de Reddit, aunque conocía los elementos básicos de la plataforma, en la que uno encontraba enlaces publicados por los usuarios que los demás usuarios votaban con una flechita hacia arriba si les gustaban o una flechita hacia abajo si no les gustaban. Los enlaces con más votos positivos aparecían en la parte superior de la página, donde los verían millones de personas. La sección de comentarios vinculada a cada publicación operaba con esas

mismas reglas. Si alguien se metía dentro de una conversación, veía primero las afirmaciones que gustaban a la multitud y ningún comentario impopular. La simplicidad y el permanente desplazamiento página abajo terminaron por atraer grandes cantidades de personas que navegaban por internet de forma esporádica, lo cual les inculcó una cultura de internet que hasta entonces había sido impenetrable. Reddit se convirtió en un portal que conectaba los dos mundos.

Aun así, Reddit se creó y se gestionaba siguiendo los mismos ideales que 4chan, y había absorbido a los usuarios y los tics culturales de esa plataforma. Los votos hacia arriba y hacia abajo imponían un mayoritarismo eclipsador que agudizó aún más la situación. Y a eso mismo contribuyó una dinámica similar a los «me gusta» de Facebook: los votos favorables se muestran públicamente, con lo cual se saca partido del impulso de validación de los usuarios motivado por el sociómetro. La segregación de dopamina enganchaba a los usuarios al sitio y, como en Facebook, guiaba sus acciones.

Los millones de usuarios que abrían Reddit todas las mañanas se encontraban con una oleada de comentarios y artículos que afirmaban la superioridad del usuario medio y se regodeaban en los agravios que este sufría. Era una versión de la realidad en la que siempre se reivindicaba el libertarismo tecnológico y en la que se humillaba y desacreditaba sin cesar cualquier sistema de creencias alternativo: el feminismo, el liberalismo convencional y la religión organizada. Si a esos usuarios medios les costaba encontrar su lugar en el mundo físico —se aseguraban unos a otros—, era porque eran más listos y más importantes que ese mundo que los trataba como personas indeseables. Al igual que 4chan, Reddit era un lugar —dicho con otras palabras— perfecto para el Gamergate.

«Empezó de una forma muy parecida a las demás plataformas —dijo Pao en referencia a los creadores de Reddit—: una pareja de hombres blancos se pone a crear algo que ellos quieren utilizar y que gusta a personas que son como ellos.» Pao había desarrollado una especial sensibilidad por el chovinismo de Silicon Valley: dos años antes, había demandado a Kleiner Perkins por discriminación por razón de género, tras lo cual la empresa la había despedido. Aunque terminó perdiendo en los tribunales, a raíz de su caso muchas personas de fuera de Silicon Valley —e incluso algunas de dentro—

declararon públicamente que había que prestar atención a las terribles desigualdades que había en el capitalismo tecnológico del valle. En 2016, cuatro años después de su litigio, solo un 11 % de los socios que aportaban capital riesgo en el sector tecnológico eran mujeres.¹ Un 2 % eran personas negras. Las empresas, por su parte, financiaban por abrumadora mayoría a personas que se parecían a ellas: en 2018, un 98 % del dinero que invirtieron estuvo destinado a empresas dirigidas por hombres.² Pao había explicado que eso reflejaba un sesgo aún más específico entre los inversores tecnológicos, un sesgo que favorecía no solo a los hombres blancos, sino también a los contestatarios idealistas de ideología libertarista; la «identificación del patrón» por la que abogaba Doerr, quien también había sido mentor de Pao. Su caso había puesto de manifiesto que las empresas que escogen a los ganadores y perdedores de Silicon Valley imponen ese arquetipo con tanta devoción porque ellas también lo encarnan.

Los directivos de Reddit, jóvenes y progresistas, no es que fueran hostiles a tales preocupaciones, al contrario. Pero la compañía convirtió en un elemento de orgullo la defensa del supuesto valor supremo de internet —una libertad de expresión total y absoluta— hasta unos extremos que otras empresas no tolerarían. Hasta que la CNN no estuvo informando durante varias semanas en tono de crítica sobre ese asunto, Reddit no cerró una subsección llamada «jailbait»,³ donde los usuarios colgaban fotos íntimas de niñas menores de edad. (Las subsecciones del portal —los subreddits— las crean y gestionan los usuarios.) Antes de cerrar la subsección, *jailbait* había sido el segundo término de búsqueda, después de *reddit*, que más usuarios traía a la plataforma.

Por Reddit circulaba casi cualquier cosa. En uno de los varios subreddits dedicados al extremismo racista, «WatchNiggersDie» («Ver a negros morir»), los usuarios colgaban vídeos atroces de personas negras muriendo en asesinatos o accidentes violentos, junto con comentarios con seudónimo como: «Casi me siento mal por dejar que una imagen como esa me llene de una alegría abrumadora. Casi...». La mayor parte de los usuarios nunca llegaban a esas comunidades, escondidas en rincones oscuros del sitio, pero accedían a ellos suficientes personas para que las organizaciones de control digital alertasen a Reddit de que se estaba convirtiendo en una incubadora de odio.

«No vamos a prohibir contenido legal, aunque nos parezca detestable o aunque lo condenemos a título personal»,⁴ había dicho Wong, director ejecutivo de la compañía. Pero su idealismo tecnológico se derrumbó en septiembre de 2014, cuando un *hacker* se infiltró en las cuentas de iCloud de varias mujeres famosas. iCloud es el servicio de computación en la nube de Apple; muchos productos de Apple hacen copias de seguridad de los archivos de los usuarios en los servidores de la empresa. El *hacker* descargó los datos del iPhone de sus víctimas, entre los cuales había varias fotos privadas en que aquellas personas aparecían desnudas. Muchas de esas fotos fueron colgadas en 4chan y luego en Reddit, que, desempeñando su papel de enlace entre los recovecos de internet y sus espacios mayoritarios, se convirtió de la noche a la mañana en el almacén principal. El sitio se vio desbordado por millones de visitantes.

«Todo el mundo hablaba de eso», dijo Pao. En las noticias, en Silicon Valley, en su vida personal. Y eso fue, en última instancia, lo que marcó la diferencia. No fue la creciente presión judicial y de la sociedad, ni fueron las víctimas, que suplicaban a las redes sociales que no permitieran esa violación de sus vidas y sus cuerpos. Eso había funcionado en todas las demás grandes plataformas, donde las fotos se eliminaron. Pero Reddit seguía teniéndolas en la plataforma. Wong convocó una reunión de toda la plantilla. Y por una vez, buena parte de los trabajadores de Reddit se unieron en contra de la comunidad del portal, que seguía insistiendo en que no se eliminasen las fotos. «La gente decía: “No sé cómo defender eso”», recordaba Pao. Y parafraseó la opinión de los empleados así: «Mis amigos y familiares me preguntan por el tema y yo no sé qué decir porque te deja mala conciencia».

Aquella polémica despertó un nivel de atención que ni siquiera el «jailbait» ni otros subreddits de odio habían despertado. El mundo exterior había tratado esos casos como transgresiones de una cultura de internet que se creía que estaba aislada del resto de nosotros. Pero eso enfrentó a Reddit y su ética *hacker* de «todo vale» con los derechos de actrices, modelos y deportistas queridas. Y el Gamergate, aunque solo estaba empezando, ya estaba apartando a mujeres inocentes de la vida pública. «¿Queremos ser como 4chan, donde básicamente hay cosas negativas y apenas puedes encontrar nada positivo? —preguntó Pao en la reunión—. Porque eso es lo que parece.» La conversación,

dijo Pao, «puso sobre la mesa una idea: “A ver, quizás esa idea de no intervenir no siempre sea la correcta”».

Pero Wong se resistía. En una publicación anunció que, a pesar de que los directivos de Reddit «entienden los perjuicios que un mal uso de nuestro sitio provoca a las víctimas de este robo»,⁵ no cederían. Reddit, escribía Wong, no era solo una plataforma social, «sino el Gobierno de un nuevo tipo de comunidad». Como muchos otros operadores de plataformas que vendrían, no obstante, dejó claro que el suyo era un Gobierno que se negaba a gobernar, pues dejaba a los propios usuarios que «eligiesen entre lo correcto y lo incorrecto, entre el bien y el mal». Tituló su publicación «Cada hombre es responsable de su propia alma». Ese se convertiría en un argumento habitual de los dueños y señores de las redes sociales: que la trascendencia de su revolución les obligaba a ignorar las insignificantes leyes y la moral del mundo físico, pasado de moda. Además, todo mal comportamiento era culpa de los usuarios, sin tener en cuenta la importancia que la plataforma hubiese tenido al permitir y fomentar esas transgresiones y por beneficiarse de ellas.

El mismo día en que Wong publicó su comunicado, unos periodistas señalaron que entre las imágenes robadas que podían encontrarse en Reddit había fotos de chicas famosas que eran menores de edad en el momento en que se habían sacado las instantáneas. Mientras tanto, los ingenieros del sitio alertaron de que, a medida que el tráfico aumentaba, tenían dificultades para mantener la plataforma en funcionamiento. Al fin, casi tres semanas después de la aparición de las fotos, Wong las eliminó. Los usuarios de Reddit, enfurecidos, acusaron a la plataforma de traicionar sus principios en favor de una influencia empresarial sombría y, peor aún, de las feministas.

Las siguientes semanas fueron un periodo convulso. Al presentar su decisión inicial como una elección ideológica, Wong había provocado que se analizasen con lupa los ideales de Silicon Valley que habían recibido poca atención a escala nacional más allá de los esmerados anuncios de Apple. El discurso según el cual estaban creando una nueva sociedad ilustrada parecía difícil de encajar con el hecho real de estar obteniendo beneficios con la explotación pornográfica de unas mujeres sin su consentimiento. En paralelo, el Gamergate estaba empeorando, y a menudo se citaba Reddit como su centro de actividad.

El consejo de administración de Reddit rechazó un plan que Wong había propuesto para reducir las oficinas. La decisión, aunque era menor, señalaba que había perdido la confianza del consejo. Entonces dimitió. En noviembre de 2014, el consejo eligió a Pao para sustituirlo. Aunque su pleito contra Kleiner seguía en el aire, el historial de Pao hacía pensar que aportaría supervisión adulta y crecimiento financiero. Pero los miembros del consejo de administración sentían tal incertidumbre que también trajeron de vuelta a uno de los fundadores del sitio, que se había ido para crear otras empresas. A él le otorgaron un cargo ejecutivo impreciso y nombraron a Pao directora ejecutiva provisional.

Sin embargo, Pao vio una oportunidad para hacer mejoras en internet que beneficiaran a todo el mundo, no solo a los hombres jóvenes blancos que dominaban las redes sociales. Describió los fiascos de Reddit como «una llamada de alerta» que indicaba la necesidad de gobernar por fin la web social. Imponiendo protecciones para las mujeres y las minorías y haciendo limpieza en subculturas de toxicidad y acoso, traería a la comunidad de Reddit lo que no había logrado incorporar en el capitalismo tecnológico basado en el riesgo: una inclusividad real.

Ese elemento pendiente del sueño de la tecnoliberación siempre había quedado en el punto ciego de la élite de hombres *geeks* de Silicon Valley. Pero Pao estaba capacitada para llevarlo a cabo. Después de un curso intensivo en las herramientas de moderación que se utilizan para gestionar la plataforma, tomó las riendas de lo que entonces era uno de los portales más visitados de internet, dispuesta a guiar a sus decenas de millones de usuarios hacia su concepción de la tierra prometida.

LOS TROLES VAN GANANDO

Cuando llevaba tres meses en el cargo, Ellen Pao se convirtió en la primera jefa de una gran plataforma que intentaba reformar los excesos de su sector. «Para mí, estaba claro que tenían que cambiar las cosas y que el acoso era excesivo», dijo. Episodios como el Gamergate, cuyo alcance ya no se podía ignorar seis meses después de su inicio, la convencieron de que había la necesidad de limitar no solo el peor

contenido, sino los incentivos y las subculturas que convertían internet en un sitio tóxico.

Pao empezó con poca cosa, prohibiendo fotos de desnudos colgadas sin el consentimiento de la persona. Quería frenar el «porno vengativo», un método para humillar a alguien divulgando fotos privadas. Las víctimas eran en una abrumadora mayoría mujeres, atacadas por exnovios furiosos o campañas de acoso al estilo de 4chan. Entendiendo a su público, Pao anunció la política como una forma de proteger no a las mujeres, sino la privacidad de los usuarios, puesto que, desde una perspectiva hipotética, los usuarios también podían ser víctimas.⁶A pesar de algunas quejas, la mayoría de los usuarios aceptaron el cambio.

Envalentonada, Pao anunció una batería de políticas que vetaban a los usuarios y a las comunidades que llevasen a cabo campañas de odio o acoso extremo. Esas iniciativas entraron en vigor de inmediato. El cambio fue, en lo superficial, modesto, e iba dirigido solo contra las conductas más indefendibles. Pero el cambio cultural que implicó fue radical, pues reorientó por primera vez los valores de gestión de una gran plataforma. Se pasaba de permitir el ello colectivo de la comunidad a controlarlo. «Fue como si mandásemos un mensaje —dijo Pao—: “No aceptamos ese comportamiento. Vamos a reprimirlo ahora y todas las veces que intentéis repetirlo”».

Pao también estaba poniendo a prueba una teoría: que las voces que propagaban más odio, aunque fueran pocas, aprovechaban la tendencia de las redes sociales a amplificar el contenido extremo gracias a su capacidad para captar la atención, lo cual teñía toda la plataforma. Acabar con «esos subreddits centrales y con altas dosis de acoso» e impedir que volvieran a aparecer, creía Pao, era la única manera segura de poner fin al «efecto dominó» del mal comportamiento. De todos modos, eso no era más que una teoría, y no estaban seguros de que funcionase.

La primera prohibición fue pequeña: un subreddit llamado «FatPeopleHate» («Odio contra las personas gordas»). Los varios miles de usuarios que tenía colgaban fotos de personas normales y corrientes que según ellos tenían sobrepeso y las convertían en objeto de acosos y amenazas. La intimidación como forma de buscar emociones fuertes sin levantarse del sofá. Pao también eliminó varios subreddits diminutos organizados en torno al odio a los negros o a las

personas LGBT. Pero los directivos de Reddit hacían hincapié en que eso se hacía por la conducta de acoso de los usuarios y no por el contenido en sí. No se eliminaron comunidades de odio mucho mayores, como «WatchNiggersDie», «GasTheKikes» («Gasear a los judíos») o «RapingWomen» («Violando mujeres»), las cuales glorificaban sin tapujos la violencia. Keegan Hankes, investigador del Southern Poverty Law Center, que monitoriza la extrema derecha, había calificado Reddit como «un agujero negro de racismo violento peor que Stormfront»,⁷ un destacado sitio web neonazi, y alertó de que incluso los peores foros supremacistas blancos prohibían tales extremos.⁸

No obstante, los usuarios de Reddit montaron en cólera por las eliminaciones, pues las consideraban un ataque a la libertad de ofender y de transgredir, que, al fin y al cabo, había sido una promesa explícita de la web social desde su creación. Como con el Gamergate, miles de usuarios votaron a favor de las publicaciones más furibundas, con lo cual fueron arrastrándose unos a otros hacia interpretaciones cada vez más extremas: Pao estaba imponiendo un programa feminista, reprimía a los hombres blancos y quería acabar con la libertad de expresión. Como ocurrió también con el Gamergate, los usuarios parecían experimentar el incidente como un ataque a su identidad compartida, un nuevo frente en la guerra imaginaria entre los hombres *geeks* librepensadores y las feministas políticamente correctas que conspiraban para controlarlos.

La plataforma recibió una avalancha de publicaciones en que se llamaba a Pao «hija de puta nazi» o «Pao Zedong», o en que se la retrataba en caricaturas de un racismo explícito, que recibieron votos favorables de forma abrumadora en una de las redes sociales más grandes del mundo. Algunos publicaban lo que pensaban que era la dirección de su domicilio, a veces acompañado de fantasías de que la pudiesen asesinar o violar allí. Pero, salvo una publicación en que se ofrecían mil dólares por una foto de Pao siendo agredida sexualmente, Pao dio instrucciones a los moderadores de no eliminar nada.

Esa serie de episodios —el Gamergate, los foros de odio, la respuesta negativa ante Pao— sacudió la web social, pero, aunque parezca extraño, no a sus supervisores, quienes no dieron señales de darse cuenta de hasta qué punto la cultura del extremismo y el mayoritarismo agresivo habían impregnado la web social. Pero no se

le pasó por alto a la extrema derecha, que vio que las redes sociales habían cultivado algo que ellos no habían sabido cultivar: un público amplio y predispuesto al nacionalismo blanco entre la juventud estadounidense. «Parece que el Gamergate puso en alerta a troles de internet racistas, misóginos y homófobos para que alcanzasen el grado de poder de que disponen en la actualidad. Lo cual es, sin duda, algo positivo»,⁹ escribió Andrew Anglin, un veterano miembro de 4chan, en el *Daily Stormer*, un destacado foro neonazi que había creado en 2013. Animó a sus seguidores a apropiarse del Gamergate y de la web social más en general, lo cual dio pie al «ascenso del ejército de troles nazis».

A su vez, otras estrellas de la web social ya estaban participando. Fredrick Brennan, creador de 8chan, había escrito un artículo de opinión en el *Daily Stormer* en favor de la eugenesia. Andrew Auernheimer, el destacado *hacker*, trabajaba para el portal como administrador del sitio web. Trabajadores de Reddit y 4chan, algunos de los cuales eran «demasiado extremos»¹⁰ incluso para los recalcitrantes supremacistas blancos que había en los foros neonazis, acudieron en tromba. El crecimiento de sus cifras pronto se vio reflejado en las plataformas convencionales, que la mayoría también usaban. Más adelante, en un estudio se calculó que el número de nacionalistas blancos de extrema derecha en Twitter se había incrementado por siete entre los años 2012 y 2016.¹¹

Pao nunca tuvo la oportunidad de llevar más lejos sus reformas. En julio de 2015, después de varios meses en que la revuelta contra ella seguía en marcha, un empleado de Reddit muy apreciado que dirigía un servicio en el que los usuarios podían hacer preguntas a personajes famosos fue despedido sin motivo. Muchas de las subsecciones más grandes del sitio, gestionadas por usuarios voluntarios, detuvieron su actividad en señal de protesta. Aunque Pao publicó un comunicado de disculpa, circuló una petición para que le rescindiesen el contrato que alcanzó las 200.000 firmas. Cuatro días más tarde, ella y el consejo de administración acordaron que dimitiría.

«Los troles van ganando»,¹² escribió Pao en un artículo de opinión en el *Washington Post* pasados unos días. A pesar de su nobleza, los ideales fundacionales de internet —alertaba Pao— habían llevado a las compañías tecnológicas a adoptar una interpretación restrictiva y extrema de la libertad de expresión que resultaba ser peligrosa.

Había durado solo ocho meses.

A pesar de todo, durante ese periodo había puesto al descubierto la encrucijada en la que se encontraban las redes sociales. Podían seguir yendo a la deriva y pareciéndose cada vez más a nuevas versiones turbo de 4chan impulsadas por algoritmos. O podían emprender un camino futuro con restricciones y normas que controlasen los impulsos de la mayoría —o de minorías especialmente vehementes—, de modo que los demás también pudiesen participar. Si bien es difícil sostener que Pao lograra convencer a su consejo de administración, a sus empleados o a sus usuarios de optar por lo segundo, sí que lo intentó. Y lo hizo casi dos años —y unas trascendentales elecciones— antes de que las críticas a escala internacional forzasen a los Mark Zuckerberg y Jack Dorsey de Silicon Valley —mucho más reticentes, tercos y poderosos— a como mínimo fingir que hacían lo mismo.

LA MAGIA DE LOS MEMES

Milo Yiannopoulos, un joven de veintinueve años que no había terminado los estudios universitarios, tenía pocos motivos para creer que alguna vez rebasaría los límites más lejanos del mundo tecnológico. Su blog de chismorreos poco fundamentado le había granjeado la posición 98 en la lista de la revista *Wired* de los cien *influencers* de 2012 («La mosca cojonera del mundo tecnológico sigue provocando e irritando») hasta que lo vendió al tener problemas jurídicos. En lo que parecía su última estación en el camino hacia la oscuridad, enviaba columnas cortas para la sección de tecnología —que casi nadie leía— de *Breitbart*, el portal web nacionalista blanco que se quedaba con la morralla. Sus primeros artículos, un cajón de sastre de chismorreos tecnológicos aburridos y quejas de la derecha sacadas de las redes sociales, no causaban gran sensación.

Sin embargo, unos meses después de empezar ese trabajo, llegó el Gamergate. Al ser un habitual de las mismas plataformas de las que había salido, Yiannopoulos sabía cómo aprovechar sus quejas y motivos, lo cual iba en la línea del programa de extrema derecha y la actitud conspirativa del portal que lo empleaba. En esas plataformas, titulares como «Las feministas intimidadoras, mentirosas y codiciosas están destruyendo el sector de los videojuegos» se hicieron virales al

interpretarse como una aparente confirmación.

Sus jefes tenían la esperanza de que sus artículos informarían a los lectores de *Breitbart*, que eran pocos y de extrema derecha, sobre asuntos tecnológicos. En lugar de eso, los textos llegaron a un público nuevo y mucho más amplio que ni siquiera sabían que existía: un público que en ese momento apenas se estaba uniendo. «Cada vez que escribes uno de tus textos, recibe diez mil comentarios»,¹³le dijo a Yiannopoulos Steve Bannon, director de *Breitbart*, en el programa de radio de la web. «Eso va incluso más allá del público de *Breitbart*; llega a todas partes.»

En tres años, la pequeña subcultura furiosa que Yiannopoulos defendía evolucionaría hasta convertirse en un movimiento mayoritario tan poderoso que le invitaron a pronunciar un discurso inaugural en el Congreso Político de Acción Conservadora, el acto más importante de la derecha política. (La invitación fue revocada más adelante.) Bannon llamó a su causa «Alt Right» ('derecha alternativa'), un término que tomó prestado de los extremistas del supremacismo blanco que querían renovar su imagen para una nueva generación. Pero para Yiannopoulos era lo que las redes sociales dijese que tenía que ser.

Yiannopoulos, al que le gusta llamar la atención con insolencia, absorbía los mensajes que parecían dominantes en internet, los exageraba un poco y volvía a publicarlos en esas mismas plataformas. Era un algoritmo viviente de las redes sociales que también podía aparecer en tertulias de televisión por cable. Se mantenía en lo más alto de los hilos de usuarios con insultos al estilo de 4chan y chistes perversos («Si la cultura de la violación fuese real, pasaría aún más tiempo que ahora en los campus universitarios estadounidenses»),¹⁴esta vez para las masas de Twitter, Facebook y YouTube. Y como nuevo rostro público de un movimiento que durante mucho tiempo no había tenido cara, provocaba o dirigía campañas de acoso e insultos colectivos.

Bannon y otros personajes de la derecha alternativa vieron una oportunidad de abrirse camino por fin. «Me di cuenta de que Milo podía conectar con esos chavales de inmediato —dijo más adelante Bannon—. Tú puedes activar ese ejército. Llegan a través del Gamergate o de lo que sea y luego se les atrae hasta la política y Trump.»¹⁵Aplicando esta teoría, a principios de 2016, *Breitbart* pidió a

Yiannopoulos que escribiera una *Guía de la derecha alternativa* de la extensión de una guía telefónica. El columnista del Gamergate, que no era un supremacista blanco de la vieja escuela ni un conspiracionista, sería el abanderado del movimiento.

Y ¿a quién mejor podía reclutar Yiannopoulos que a Andrew Auernheimer, el famoso *hacker* que se había tatuado una esvástica en el pecho? «Por fin estoy redactando mi gran artículo sobre la derecha alternativa —escribió Yiannopoulos en un correo a Auernheimer—. ¿Te apetece pasarme algunas ideas?»¹⁶ Tras hablar por correo con otros ultranacionalistas y personajes supremacistas blancos, Yiannopoulos envió su material a un redactor profesional que había contratado a través del Gamergate y le pidió que diera unidad a los textos para armar un artículo que aparecería bajo su nombre. «La derecha alternativa es un movimiento nacido de los márgenes jóvenes, subversivos y alternativos de internet»,¹⁷ se decía. Los habitantes de la web serían los propagandistas de la derecha alternativa, quienes llevarían a un público más general una causa que «promete diversión, transgresión y un desafío a las normas sociales que ellos, sencillamente, no entienden».

Semana tras semana, el lenguaje de los extremistas de la web social fue filtrándose a las grandes plataformas, lo cual lo hizo llegar al ámbito político general. Durante las semanas iniciales de las primarias presidenciales republicanas, las plataformas se llenaron de acusaciones de que los candidatos moderados eran unos *cuckservatives* ('conservadores cornudos'). Se cree que el compuesto se originó en foros supremacistas blancos, donde *cuck* ('cornudo') hace referencia al miedo a que tu esposa te ponga los cuernos con un negro.¹⁸ Los *gamergaters* trajeron el término a Twitter, donde por su llamativa procacidad llegó a los primeros puestos de los hilos de los usuarios. Los conservadores mayoritarios fueron adoptándolo de un modo paulatino, con lo cual se introdujo en el espacio de la derecha —que se disputaba cada vez más en Twitter y Facebook— una influencia de la cultura del troleo que no dejaría de crecer.

«Lo llaman “la magia de los memes”: cuando memes de internet que antes eran desconocidos se vuelven tan influyentes que empiezan a afectar a los hechos del mundo real»,¹⁹ escribió Yiannopoulos ese verano antes de las elecciones. Una gran cantidad de memes de la derecha alternativa siguieron un camino parecido. La rana Pepe, un

personaje que los neonazis de internet habían adoptado como mascota, se convirtió, en Twitter y Reddit, en algo más complejo: una expresión de racismo irónico. Los memes de Pepe deliberadamente ofensivos (una rana Pepe nazi manejando una cámara de gas, una rana Pepe judía sonriendo por los atentados del 11 de septiembre) se volvieron la peineta que se le hacía al mundo. Y era un vehículo para introducir a hurtadillas, escondido detrás de la ironía, un racismo terriblemente real en el discurso mayoritario. «Los no adoctrinados deberían ser capaces de entender si estamos bromeando o no —escribió Anglin, el fundador del *Daily Stormer*, en la guía de estilo y contratación del sitio web—. Esto es, obviamente, una estrategia y, en realidad, sí quiero gasear a los judíos. Pero eso no aparece por ninguna parte.»²⁰

El movimiento convergió alrededor de Donald Trump, quien se había consolidado con los mismos tics y tácticas que Yiannopoulos y otras estrellas del Gamergate y, al parecer, por la misma razón: era lo que premiaban las redes sociales. Blandía como armas la desinformación y la misoginia. Troleaba sin vergüenza, acumulando víctimas con burlas e insultos. Desafiaba a los guardianes de la sociedad a que se ofendieran por provocaciones extravagantes salidas de 4chan. «Tiene un carácter y un estilo que encajan por completo con lo que buscan los malhechores de la web»,²¹ escribió Yiannopoulos. En una popular versión de la rana Pepe, el animal llevaba la corbata roja y el pelo rubio de Trump. Trump publicó la imagen en Twitter, y su hijo en Instagram.

A pesar de todo, aunque Yiannopoulos y Bannon llevan mucho tiempo afirmando que son ellos quienes fusionaron la cultura del troleo en internet con la derecha generalista, en realidad el mérito recae en una fuerza mucho más poderosa: Facebook. Según se descubrió en un estudio de Harvard,²² desde mayo de 2015 —un mes antes de que Trump hiciera oficial su candidatura— hasta noviembre de 2016, la fuente de información de derechas más popular en Facebook fue *Breitbart*, que dejó atrás incluso a Fox News. Fue el tercer medio de comunicación con más comparticiones de todos, superando a todos los demás periódicos y cadenas de televisión salvo la CNN y el *New York Times*. Si alguien leía noticias con regularidad en Facebook en 2016 —lo cual ese año hicieron un 43 % de los estadounidenses—,²³ probablemente leía *Breitbart*. Llegó a ser tan

dominante en la plataforma que, incluso a finales de 2019, después de que el portal hubiese perdido popularidad a raíz de una mala gestión y de varias polémicas, Facebook designó *Breitbart* como una «fuente de noticias fiable»²⁴ con un acceso especial a los lectores de Facebook.

Durante mucho tiempo, los que desconocían ese mundo adscribieron asombrados el ascenso de *Breitbart* a una manipulación de las redes sociales con artes oscuras. Pero, en realidad, la publicación hizo poco más que publicar sus artículos en Facebook y Twitter, como siempre había hecho. Fue, en muchos sentidos, un beneficiario pasivo. Los sistemas de Facebook promovían un buen número de blogs ultrapartidistas antes desconocidos y medios que vendían pura desinformación —con nombres como *The Gateway Pundit*, *Infowars*, *The Conservative Treehouse* o *Young Cons*— hasta convertirlos en megaeditores con el poder de reconfigurar la realidad para enormes segmentos de la población.

Como sucedió con el Gamergate, no era solo que el contenido fuera lascivo. Promovía sentimientos de amenaza a la identidad de grupo —los musulmanes vienen a matarnos, los judíos arrasarán nuestra cultura, los liberales quieren acabar con los valores tradicionales— que activaban el instinto de defensa tribal de los usuarios que Henri Tajfel y su equipo de psicólogos sociales habían identificado décadas antes. Los incitaba a compartir cada vez más enlaces y comentarios que consolidaban su identidad de endogrupo y a unirse contra el enemigo común. Eso les hacía seguir clicando y publicando, lo cual espoleaba a otros usuarios a caer en el mismo frenesí, un bucle de retroalimentación continuo de miedo y rabia que resultó ser de lo más beneficioso para Silicon Valley y Donald Trump, pero catastrófico para todo el resto.

El alcance y el algoritmo de Facebook concedieron a la compañía una importancia infrecuente en la creación de ese ecosistema informativo. Pero en las demás plataformas tuvo lugar más o menos el mismo proceso, lo cual hacía pensar que el problema no se debía a alguna peculiaridad de Facebook o de sus creadores, sino que era algo endémico de las redes sociales modernas. En Twitter, *Breitbart* se convirtió en «el nexo de los medios conservadores» y la fuente de información más compartida entre los seguidores de Trump, según concluía el estudio de Harvard.²⁵ *Breitbart* dominaba, en concreto, los debates acerca de la inmigración, quizás la cuestión más polémica de

la campaña presidencial de 2016. Sus artículos sobre inmigración se compartieron, en conjunto, más del doble que los de cualquier otro medio de comunicación, una cifra espectacular teniendo en cuenta que solo contaba con unos pocos periodistas.

En todas las plataformas, los debates sobre la inmigración «giraron con mayor frecuencia en torno a cuestiones relacionadas con la amenaza a la identidad», concluyeron los autores del estudio, lo cual, como es lógico, privilegió las visiones de extrema derecha que se centran, por definición, en los temores de un conflicto identitario. El estudio llegó a la conclusión de que los «medios muy partidistas» que divulgaban desinformación y los «portales políticos ultrapartidistas basados en el ciberanzuelo e impulsados por Facebook» eran endémicos en 2016, una tendencia que «tuvo una importancia mucho mayor en la derecha que en la izquierda».

Si hubiéramos querido saber adónde nos estaba llevando aquello, lo único que tendríamos que haber hecho habría sido fijarnos en el Gamergate, que, en verano de 2016, había declarado más o menos la victoria en su guerra cultural por los videojuegos. «La consecuencia a largo plazo del Gamergate fue la creación de una cultura del miedo dirigida contra las mujeres profesionales que hacen tus videojuegos», dijo Brianna Wu. Muchas dejaron el sector derrotadas. Las que se quedaron aprendieron las consecuencias de levantar la voz. El sector, que genera cientos de miles de millones de dólares, todavía tiene un sesgo —tanto en sus productos como en las personas que los crean— radicalmente masculino.

Como la mayoría de los observadores —le conté a Wu—, yo al principio no consideré que el Gamergate fuera un presagio del mundo que vendría y en ese momento no tuve en cuenta sus consecuencias. «Si te digo la verdad —dijo ella—, yo también lo subestimé. No podía prever que nuestra política en su conjunto se convertiría en algo como el Gamergate. Y en efecto, así fue.» Agregó que le gustaría haber previsto «el número de personas que están empleando tácticas al estilo del Gamergate. Seamos claros: es sobre todo la derecha quien lo está haciendo. Pero también hay muchas personas en la izquierda que he visto con horror que utilizaban la misma táctica de fomentar la indignación y esa cultura de turba y humillación».

El resultado fue una convergencia casi universal en esos comportamientos y formas de pensar, incentivados desde el principio

por las redes sociales. Hubo momentos, admitió Wu, en los que se vio tentada a fomentar la indignación, a unir a sus seguidores en contra de algún adversario, a difundir una afirmación que, aunque fuera dudosa, podía fortalecer la identidad y acrecentar los prejuicios de su endogrupo. La mayoría de las veces se contuvo, pero no todas.

«Es algo que me cuesta, a mí personalmente, de la manera en la que interactúo con el mundo, porque con todo esto se ha desatado algo muy peligroso —confesó—. Ese ciclo de agravios y resentimiento e identidad, y de muchedumbres enfurecidas, parece que está consumiendo y envenenando a todo el país.»

La tiranía de los primos

JUSTICIA VIRTUAL

En septiembre de 2015, apenas dos meses después de que Ellen Pao fuera destituida de Reddit, Walter Palmer, un dentista de cincuenta y cinco años de un barrio residencial de Mineápolis, se reunió con dos periodistas de un periódico para insistir en que no se estaba escondiendo. «He estado con gente, familia y amigos. La ubicación no es en realidad tan importante —dijo Palmer, flanqueado por un abogado y un representante de relaciones públicas—. Al principio, mi hija y mi esposa tuvieron algunos problemas de seguridad.»¹ En lo que dijo que sería su única entrevista, Palmer criticó el movimiento internacional —el primero de una nueva clase de rabia masiva y alteradora de la vida que pronto se convertiría en algo habitual— que acosaba a su familia y empleados, quienes, según afirmaba él, eran inocentes. «No entiendo ese nivel de inhumanidad para perseguir a unas personas que no tienen ninguna implicación»,² dijo.

Seis semanas antes, la BBC había retransmitido una pequeña pieza informando de que un cazador no identificado había matado un león en Zimbabue.³ Eso no era inhabitual ni ilegal de por sí. Pero los guías del cazador habían atraído el león hasta el exterior de una reserva de caza cercana, donde al parecer sí estaba prohibido matar animales. Y el animal tenía una distintiva melena negra que lo había hecho famoso entre los visitantes de la reserva, que le habían puesto el nombre de Cecil.

Una de las subsecciones más populares de Reddit en esa época era la de noticias internacionales, un repositorio de informaciones del mundo y de chismes de agencias de noticias. Como los usuarios colgaban más de un millar de enlaces cada día, alcanzar los puestos más elevados de la página requería una descripción diseñada para provocar una emoción intensa. Un usuario con el alias Fuckaduck22 —el que publicó la noticia sobre Cecil— agregó este título: «El león

Cecil fue desollado y decapitado / Hwange. Un león de melena negra muy conocido y muy fotografiado llamado cariñosamente Cecil murió a manos de unos cazadores deportivos a poca distancia de Hwange, en Zimbabue, la semana pasada».

Miles de usuarios votaron a favor del artículo, que llegó a la parte superior de la página de inicio de Reddit, lo cual le dio un público de millones de personas. No todas estaban enfadadas. A fin de cuentas, Cecil estaba cerca de la esperanza de vida de los leones machos, señaló un usuario, mientras que otro indicó que las expediciones de caza mayor financian proyectos de conservación. Pero las secciones de comentarios de Reddit están clasificadas según la popularidad de cada comentario. Y lo que llegó a los puestos más altos fueron las expresiones de rabia:

los leones en estado salvaje podrían extinguirse pronto porque a ti te apetecía UN MONTÓN hinchar tu masculinidad en declive.

Yo quiero la cabeza de ese cazador como trofeo. Puto cobarde de mierda.

Siguiendo de un modo inconsciente los incentivos de la plataforma, los usuarios la llenaron de comentarios sobre Cecil, y en cada ronda se incrementaba la implicación emocional. Primero, la muerte del león era una triste pérdida; luego, un golpe devastador para los amantes de los animales de todo el planeta, y, por último, un crimen indignante. Primero, el cazador era un tremendo cobarde; luego, un sangriento asesino, y, por último, se dijo que estaba «afectado por una enfermedad mental o alguna forma de psicopatía».

En verano de 2015, periodistas y editores aún estaban aprendiendo a seguir la viralidad de las redes sociales, lo que podía multiplicar su público varias veces. Eso fue lo que ocurrió cuando usuarios de Reddit llevaron a Cecil hasta el primer puesto de la lista de los más leídos en muchos medios de comunicación: investigaron a fondo un incidente que de otro modo habrían ignorado. Se hizo público que el cazador, hasta entonces anónimo, era Palmer, el dentista. Aun así, la noticia parecía inofensiva. Hasta que saltó a Twitter.

Los creadores de Twitter habían puesto en marcha su producto en

2006, que en esencia era un servicio de mensajes de texto en grupo. Los usuarios mandaban un mensaje de texto a un número gratuito, que luego transmitía el mensaje a sus amigos.⁴ Si salías a tomar algo, podías publicar el nombre del bar, y los amigos que seguían tus actualizaciones podían sumarse a la fiesta. Los mensajes estaban limitados más o menos a la longitud de una frase. Una página web sencilla también registraba las actualizaciones. Jack Dorsey, en una prueba inicial en la que imaginaba cómo funcionaría el servicio, escribió mensajes de muestra como «en la cama» o «de camino al parque».⁵

Por defecto, las publicaciones eran públicas; ¿por qué esconder unas actualizaciones anodinas detrás de un engorroso muro de privacidad? Ese grado de apertura ganó importancia en 2009, cuando, en un país remoto como Irán, los ciudadanos llenaron las calles para protestar por unas elecciones que creían que el Estado había manipulado. Cuando las autoridades tomaron medidas represivas, algunos manifestantes usaron Twitter para escribir a toda velocidad actualizaciones sobre las barricadas. Cuando se dieron cuenta de que el mundo los estaba mirando —y que la observación internacional podía disuadir o por lo menos documentar los abusos gubernamentales— sus tuits cobraron mayor importancia. En medio del optimismo tecnológico de la era de Obama, los estadounidenses lo describieron como una «revolución nacida en Twitter».⁶ Aunque la revolución real de los iraníes fue reprimida con rapidez, la etiqueta permaneció.

A medida que otros eventos internacionales se desarrollaron en directo en Twitter, cualquier persona a la que le importaban las noticias o que se dedicaba a crear información se unió. Evolucionó de un servicio de mensajes en grupo a una plataforma de publicaciones cortas de cara al público: un microblog donde millones de personas retransmitían y debatían de forma colectiva las maquinaciones políticas del día o un episodio de *Juego de tronos*. Pero todo se filtraba a través de una plataforma que opera más como Reddit que como un teletipo de la CNN. Los usuarios validan las publicaciones de los demás usuarios clicando «me gusta» o haciendo retuit, con lo cual hacen subir a la superficie el contenido que más gusta a su voluntad colectiva o lo relegan a la irrelevancia.

Cuando la rabia por Cecil alcanzó los 316 millones de usuarios

activos de Twitter, el nivel de actividad ascendió incluso más allá de los extremos del Gamergate. En un solo día, los usuarios publicaron 672.000 tuits sobre el incidente.⁷ Unos 50.000 nombraban a Palmer directamente y obtuvieron un total de 222 millones de visualizaciones: una infamia tal que, por lo común, solo pueden infligir los libros de historia. «Voy a pagar más de 35.000 libras para ver a #WalterPalmer en un combate en igualdad de condiciones con el rey de la selva»,⁸ escribió un usuario, representante de una oleada de publicaciones en las que se expresaba el deseo de ver sufrir a Palmer.

Cuando se incrementaba la rabia en Reddit, esa permanecía limitada sobre todo a nichos de *geeks* de internet. Pero Twitter, ya en 2015, estaba en el centro de la actualidad y la cultura de masas. Las fantasías de venganza inundaron lo que era el equivalente en redes sociales de una cadena de televisión, lo cual atrajo a periodistas y famosos. «Espero que #WalterPalmer pierda su casa, su consulta y su dinero», tuiteó Sharon Osbourne, en ese entonces estrella de un *reality show*, mientras que la galardonada actriz Mia Farrow publicó la dirección de la clínica dental de Palmer.⁹

Cuando las plataformas se convierten en un cántico al unísono de «A por él», luego suelen pasar cosas. Cientos de personas publicaron reseñas negativas de la clínica dental de Palmer en webs como Yelp con la esperanza de hacerle perder clientes.¹⁰ En pocos días, hubo tantos tuits con amenazas violentas que su consulta cerró por completo. (Cuando un usuario de Reddit sugirió que eso podía perjudicar injustamente a los empleados de Palmer, 1.500 usuarios votaron a favor de la respuesta «A sus empleados les convendría más trabajar en cualquier otra parte».¹¹) Aparecieron pintadas amenazadoras en las puertas de la clínica dental. Alguien pintó con espray la casa de Palmer. Este y su familia, quizás preguntándose si alguna de las personas que merodeaban por los alrededores de su casa podían pasar a la acción, se fueron a lo que todo el mundo salvo Palmer describió como un escondite.

Cuando Palmer contrató a un representante de relaciones públicas y se reunió con periodistas, se habían publicado más de dos mil artículos sobre su cacería en Facebook, donde se habían compartido 3,6 millones de veces.¹² El *New York Times* publicó un editorial formal lamentando la pérdida de Cecil; el gobernador de Minesota condenó los hechos; Jimmy Kimmel, presentador de un *late*

night, se quedó sin palabras durante el programa. Todo ello provocado por la viralidad de las redes sociales. Entonces, un día, la fiebre se apagó y todo el mundo siguió con su vida.

El arrebató dividió a los analistas del mundo tecnológico. Algunos vieron otro Gamergate, ahora un poquito más cerca del centro de la sociedad estadounidense, un hecho que había sido posible gracias a unas plataformas que promovían una histeria de muchedumbre y un acoso casi sin restricciones ni amparo. Otros vieron un drama absurdo de internet sin mayor importancia y tampoco encontraron motivos para culpar a las redes sociales, que aún se consideraban elementos liberadores a escala internacional.

James Williams, ingeniero de Google desde 2005, encajaría en el perfil del sector más optimista. No obstante, al dejar la compañía, escribió un artículo para el portal *Quillette* diciendo que el episodio de Cecil y otros parecidos auguraban un cambio radical en nuestros mundos en proceso de digitalización.¹³ Los vengadores de Cecil — fueran cuales fuesen sus motivaciones— habían sido manipulados de forma inconsciente para comportarse de aquella manera, creía Williams, «los botones se los habían pulsado los patrones persuasivos del diseño digital». Ese era su diagnóstico: «Nos encanta indignarnos. Respondemos a eso como una recompensa».

Las plataformas habían aprendido a dar rienda suelta a la indignación que provocaba en sus usuarios «un subidón: de razón de ser, de claridad moral, de solidaridad social». El ritmo creciente de esos arrebatos totales y absolutos, quizás uno por semana, indicaba que las redes sociales no estaban influyendo solo en la cultura más mayoritaria, sino que hasta cierto punto la estaban sustituyendo, en beneficio de —y en ese momento eso era un argumento extravagante— Donald Trump. ¿Era tan descabellado pensar —proponía Williams— que los nuevos medios dominantes de nuestra época pudieran catapultar a abanderados como Trump hasta los cargos de mayor poder? ¿No estaba ocurriendo ya? «La verdad es que —escribía Williams— estos efectos políticos ya están aquí.»¹⁴

INDIGNACIÓN MORAL

Billy Brady era estudiante de primer año de la Universidad de

Carolina del Norte cuando se dio cuenta de que disfrutaba indignándose en Facebook. Durante todo el año 2005 convirtió las visitas a Facebook en algo habitual, batallando en las secciones de comentarios y, como ahora admite riéndose, «publicando cosas bastante incendiarias».

Brady, que es un joven vegano y tranquilo que habla con un acento de surfista californiano sin ser de California, no estaba acostumbrado a sentirse atraído por la rabia y los conflictos. Y se había creado un perfil de Facebook para convencer a otras personas, no para pelearse. Se estaba iniciando en el activismo por los derechos de los animales —me contó— y pensaba que «[Facebook] parecía una plataforma interesante para divulgar mensajes y convencer a los demás». Pero a menudo terminaba expresándose con indignación ante los otros usuarios. Se dio cuenta de que se comportaba de tal forma que tenía poca probabilidad de hacer progresar la causa y, aunque fuera divertida en ese momento, después su conducta le hacía sentirse como un imbécil.

Por suerte, la rama de estudio de Brady pudo ayudarle a entender la situación. La filosofía moral, antaño un dominio de señores que se rascaban la barbilla, se estaba volviendo más empírica, aprovechando la ciencia aplicada para desentrañar la naturaleza real de la moralidad. Las conclusiones de la filosofía moral proponían una explicación para el comportamiento de Brady. Indignarse en su escritorio o proferir un insulto le hacían sentir peor, no mejor. Pero cuando sus expresiones de rabia atraían la atención en internet, sobre todo los ánimos que le daban otras personas que pensaban como él, el subidón era adictivo. «Me di cuenta claramente de que puedes captar la atención de la gente publicando comentarios más incendiarios y sintiendo ciertas emociones», me contó.

Pero ¿por qué las personas se sentían atraídas por unas emociones tan perjudiciales y desagradables? Fascinado, se inscribió a algunas asignaturas de Psicología. Desde entonces ha estudiado la influencia de las redes sociales en la conducta. En cierto sentido, me dijo, todos los usuarios están al mismo tiempo comportándose y sirviendo como participantes de un experimento psicológico permanente. Por regla general, las personas están muy atentas y adaptadas a los comentarios sociales, un impulso que aprovechan los «me gusta» y las comparticiones digitales. Incluso cuando era un

universitario que se metía en disputas en Facebook, «por ensayo y error, aprendí cómo la gente responde a distintos marcos y a diferentes apelaciones».

Cuanto más él y los usuarios que compartían su mismo parecer habían ido retroalimentando su rabia, más hostil se había vuelto su comportamiento. Los carnívoros y los escépticos de los derechos de los animales no eran personas bien intencionadas a las que había que convencer, sino unos catetos y unos malvados a los que había que condenar. La clave, concluyó Brady, era la indignación moral. La indignación es un cóctel emocional sencillo: rabia más asco. La indignación moral es un instinto social.

¿Os acordáis de aquellas tribus primitivas de hasta 150 personas? Para sobrevivir, el grupo tenía que garantizar que todo el mundo actuase en favor del interés colectivo, y eso en parte consistía en llevarse bien unos con otros. Requería un código de conducta compartido. Pero ¿cómo se logra que todo el mundo interiorice y siga ese código? Pues la indignación moral es la adaptación de nuestras especies ante ese desafío. Cuando ves a alguien infringiendo una norma importante, te enfadas. Quieres que se le castigue. Y te sientes obligado a expresar esa rabia, de tal forma que los demás miembros del grupo también vean la infracción y quieran participar en la humillación y quizás en el castigo del transgresor.¹⁵

El deseo de castigar a los infractores es tan profundo que aparece incluso en los bebés. En una serie de experimentos, a varios niños de menos de un año les enseñaron dos marionetas.¹⁶ Un muñeco compartía; el otro se negaba a hacerlo. Los bebés sistemáticamente cogían caramelos de la marioneta mala y recompensaban a la buena. Los participantes de tan solo uno o dos años más incluso recompensaban a las marionetas que eran crueles con la marioneta mala y castigaban a las marionetas que eran amables con ella. Era una confirmación de que la indignación moral no es solo sentir rabia contra los transgresores. Es un deseo de ver como la comunidad entera se une contra ellos.

Brady —con el objetivo de entender la indignación que había sentido en internet, primero cursando un máster de Filosofía y luego otro de Psicología— dio vueltas alrededor de una teoría que estos dos campos llaman sentimentalismo. «Es la idea de que nuestro sentido de la moralidad está entrelazado con nuestras respuestas emocionales y

quizás incluso impulsado por ellas —dijo—. Lo cual contradice esa vieja idea de que los humanos son muy racionales en términos morales.»

La cultura popular suele retratar la moralidad como si emergiera de nuestros principios más íntimos: el ángel que llevamos dentro, la mente ilustrada. El sentimentalismo dice que, en realidad, la moralidad está motivada por impulsos sociales como la conformidad y la gestión de la reputación (¿os acordáis del sociómetro?), los cuales experimentamos como emociones.¹⁷ Hay investigaciones en neurología que confirman eso.¹⁸ Cuando las personas que se encuentran ante un dilema moral deciden cómo responder, presentan una intensa actividad en regiones neuronales asociadas con las emociones. Y el cerebro emocional funciona deprisa, a menudo toma decisiones antes de que la razón consciente tenga siquiera oportunidad de hacer acto de presencia. Hasta que no les pidieron que explicasen su elección, los sujetos no activaron las partes del cerebro responsables del cálculo racional, que usaron, de forma retroactiva, para justificar cualquier acción motivada por las emociones que hubieran decidido de antemano.

Esas elecciones morales-emocionales parecían estar claramente al servicio de un objetivo social, como buscar la aprobación de los iguales, recompensar a un buen samaritano o castigar a un transgresor. Pero la naturaleza instintiva de ese comportamiento permite que sea susceptible a la manipulación. Y eso es justo lo que han aprendido a hacer déspotas, extremistas y propagandistas, con lo cual pueden poner a la gente de su lado generándole indignación, a menudo contra algún chivo expiatorio o un malhechor imaginario. ¿Qué sucedería si, inevitablemente, las plataformas sociales aprendieran a hacer lo mismo?

TURBAS HUMILLADORAS

Brianna Wu aún estaba enfrentándose a una muchedumbre indignada en las redes sociales cuando intentó espolear a otra. «Nunca lo olvidaré —me dijo—. Hubo un momento, por la mitad del Gamergate, en el que yo estaba muy metida en el activismo feminista dentro del sector de los videojuegos.» Oculus, una empresa propiedad de

Facebook que fabricaba gafas de realidad virtual, había anunciado un nuevo equipo. Todos hombres blancos. Wu estaba enfurecida. «Recuerdo que exploté en Twitter; publiqué la imagen del equipo y los nombres de sus miembros —me dijo—. Y eso es un problema. Es terrible la falta de mujeres en los equipos de ingenieros que trabajan en realidad virtual. Pero me di cuenta de que esas tácticas funcionan en el mismo plano emocional que lo que el Gamergate me estaba haciendo a mí.»

Por primera vez, Wu reconoció delante de sí misma hasta qué punto su experiencia en las redes sociales orbitaba en torno a la provocación o la participación en campañas de humillación pública. Algunas duraban semanas, mientras que otras solo unos minutos. Algunas estaban al servicio de una causa importante —o eso se decía a sí misma—; otras solo porque alguien le había tocado las narices. «Solía escoger tuits sexistas que me habían dirigido. Los retuiteaba y dejaba que mis seguidores virtuales se ocupasen del asunto —dijo—. Eso ya no lo hago más porque es, básicamente, pedir que se acose a una persona.»

En 2016, Wu se presentó a las primarias del Partido Demócrata para conseguir el escaño de su distrito en la Cámara de Representantes defendiendo causas de izquierdas, junto con la promesa de abordar el acoso en internet. Un periodista le preguntó por un tuit de 2014 en el que había identificado el nombre y la empresa de alguien que había publicado algo sexista: ¿no había sido eso un acto de acoso como los que ella instaba las plataformas a combatir? «Recuerdo que en ese momento pensé: “Vaya hipócrita de mierda que soy”. No tengo excusa ante eso —dijo—. Y eso es lo que dije. “Me equivoqué”.» Había dejado que las plataformas sacasen en ella el mismo comportamiento que aborrecía en los demás, confesó. «Y no sé cómo toda esa mierda puede volverse menos tóxica si no es logrando que más personas se den cuenta de que, en nuestros peores momentos, todos podemos ser el malo de la película.»

En esa época, Kate Klonick, doctora en Derecho por la Universidad de Yale que había estudiado Psicología Cognitiva, intentaba entender por qué las humillaciones masivas parecían tan extendidas en las redes sociales. Encontró muchas claves para averiguarlo en una polémica de los primeros tiempos de Twitter conocida como el Donglegate, una denominación que demuestra la

poca seriedad con que se tomaban las cosas en aquel entonces. Una programadora llamada Adria Richards, mientras estaba sentada entre el público en un congreso de desarrollo de *software*, había oído a un hombre que estaba en la fila de detrás de ella medio susurrar un chiste sobre terminología informática al chico que tenía a su lado.¹⁹ Ella se dio la vuelta y sacó una foto, y luego la publicó en Twitter con el texto: «No mola. Por aquí están haciendo chistes sobre bifurcar repositorios con insinuaciones sexuales y “grandes” adaptadores.²⁰ Justo detrás de mí».

Richards, una mujer negra, llevaba tiempo levantando la voz contra los sesgos de su sector. Ahora, a comienzos de 2013, las redes sociales le permitieron llegar a un público mucho más amplio. Twitter y Facebook se encendieron con publicaciones enfurecidas en reacción a su tuit. De repente, los organizadores del congreso —y también las empresas de los dos hombres— eran considerados infames o incluso detestables, y eso en todo el país. Uno de los dos fue despedido al cabo de poco tiempo. Se disculpó por su comentario en *Hacker News*, el foro virtual oficioso de Silicon Valley, diciendo que lo sentía sobre todo por Richards. Pero añadía: «Tengo tres hijos y me gustaba mucho ese trabajo. No me avisó, sonrió mientras hacía la foto y condenó mi existencia. Que eso sirva como un mensaje para todo el mundo: nuestros actos y palabras, ya sean grandes o pequeños, pueden tener graves consecuencias». Inevitablemente, vino una segunda ola de indignación, que esta vez inundó a Richards y a sus jefes con mensajes furiosos y a veces amenazantes. La página web de su empresa quedó inoperativa por culpa de un truco informático rudimentario llamado denegación de servicio. Al día siguiente, Richards fue despedida por —en palabras de su jefe— haber «humillado públicamente» a aquellos dos hombres.

Klonick sabía que buena parte de los expertos en derecho consideran que las humillaciones públicas son necesarias para el funcionamiento de la sociedad: mostrar rechazo ante alguien por colarse, evitar a una persona por haber hecho un comentario sexista o despedir a alguien por unirse a una organización que fomenta el odio. Pero las redes sociales estaban alterando el funcionamiento de esas humillaciones públicas, lo cual cambiaría, por fuerza, el funcionamiento de la propia sociedad. «El acceso barato, anónimo, instantáneo y omnipresente a internet ha acabado con la mayoría de

los límites naturales (si no con todos) a las humillaciones —escribió Klonick sobre sus hallazgos—, y de este modo ha transformado la forma en que percibimos e imponemos las normas sociales.»²¹

Antes, alguien que era, por ejemplo, maleducado con el conductor del autobús podía ser objeto de comentarios por parte de algunos pasajeros. Ahora, si el incidente se graba y se cuelga a internet, esa persona podría sufrir semanas de insultos llegados desde todo el planeta. «Hoy es más fácil que nunca —escribía Klonick— utilizar la humillación para hacer cumplir las llamadas normas sociales y es más fácil que nunca que esa humillación se descontrole.» Las humillaciones públicas en el ámbito virtual solían ser «excesivas» —sostenía Klonick—, mal calibradas para la escala del delito y «con poca o una cuestionable precisión con respecto a quién y qué se castiga».²²

Además, las humillaciones se habían vuelto más crueles e incluso sádicas. Desde el momento en que la web social se enlazó con la vida cotidiana, empezaron a circular muchedumbres indignadas que se excedían. En 2013, una trabajadora llamada Justine Sacco se hizo tristemente famosa a escala internacional cuando un tuit dirigido a sus 170 seguidores («De camino a África. Espero no coger el sida. Es broma. ¡Soy blanca!») provocó decenas de miles de respuestas furiosas y luego los usuarios se regodearon en su humillación y posterior despido. Un artículo del *New York Times Magazine* sugería que a Sacco y a otras personas como ella se las había hecho sufrir para divertirnos o sencillamente porque habíamos perdido el control, «como si ahora las humillaciones se hicieran por la humillación en sí, como si siguieran un guion».²³

La foto de un aula de una pequeña ciudad de Wisconsin se hizo viral cuando los usuarios de Twitter acusaron a los alumnos de hacer la salutación nazi.²⁴ Una tormenta de rabia internacional, toda virtual, puso patas arriba las vidas de los padres y de los responsables del colegio, que fueron humillados con cada detalle suyo que los usuarios pudieron encontrar en internet, lo cual dejó cicatrices en la ciudad. Más adelante se demostró que el fotógrafo voluntario había captado a los chicos saludando desde un ángulo raro.

Un periodista novato de Des Moines escribió un artículo sobre un hombre de la zona que se había hecho viral en Facebook por asistir a un partido de fútbol americano universitario con un cartel gracioso

(«Busch Light Supply Needs Replenished», ‘Hay que rellenar las reservas de cerveza Busch Light’), seguido de su usuario de Venmo.²⁵El periodista señalaba de paso que, tiempo atrás, el hombre había publicado unos tuits ofensivos por los que ya se había disculpado. Primero, los usuarios de las redes sociales exigieron que el portador del cartel perdiese su trabajo y luego, llevando la reacción negativa al extremo, que también despidieran al periodista, lo cual sucedió.

Un hombre palestino-americano, dueño de un restaurante en Mineápolis, fue acosado por unos usuarios furiosos en internet cuando alguien descubrió, en el historial de las redes sociales de su hija, que la chica había publicado comentarios racistas cuando era pequeña. Aunque ambos se disculparon, los usuarios divulgaron su dirección y presionaron a las empresas locales para que dieran la espalda a la familia. Les rescindieron el contrato de alquiler del restaurante y dos tiendas de comestibles retiraron los productos de la familia. La indignación se apaciguó poco después de que el restaurador despidiese a su hija; y le dijo a un periodista: «¿Cree usted que es fácil para mí destruir la carrera profesional de mi hija?».²⁶

Cuando una novelista inmigrante china obtuvo un contrato muy cuantioso por su primer libro, una obra de narrativa fantástica para un público juvenil, otros escritores hicieron tuits vagos afirmando que el original contenía fragmentos racistas y plagiados. En Twitter mucha gente se tomó los rumores en serio, acosó a la escritora y a su editorial hasta que retiraron el libro. Más adelante se demostró que las acusaciones eran falsas. Un mes después, volvió a ocurrir lo mismo, esta vez contra un joven novelista negro, y se consiguió presionarlo para que también retirase el original. Incidentes como esos han sido recurrentes entre novelistas, lo cual ha hecho que un crítico literario afirmara exasperado: «Es desconcertante que a personas que se pasan el día contando con entusiasmo lo mucho que les gusta leer se les dé tan mal la lectura».²⁷

Una estudiante negra del Smith College publicó en Facebook que un conserje y un guardia de seguridad la habían acosado por «comer aun siendo negra».²⁸La habían interrogado mientras almorzaba en un bar de la residencia de estudiantes, dijo, tratándola como a una intrusa debido a su raza. Compartieron la publicación muchas personas indignadas del Smith College, luego saltó a otras

universidades y después a la sociedad en general, lo cual atrajo una atención inusitada sobre aquel minúsculo campus. Al conserje se le concedió un permiso retribuido. Los estudiantes hicieron huelga en señal de protesta. La estudiante publicó una segunda parte en Facebook acusando a dos trabajadores de la cafetería de haber llamado a seguridad durante el incidente. Incluyó el nombre, la dirección de correo y la foto de uno de ellos, y escribió: «Aquí tenéis a la persona racista». La trabajadora de la cafetería recibió montones de llamadas furiosas a su casa, algunas amenazándola de muerte. La estudiante también publicó una foto del conserje, acusándole de «actos racistas y cobardes», si bien, al parecer por error, había identificado al conserje equivocado.

La verdad resultó ser una combinación de un sincero malentendido y de una exageración juvenil. La chica se había sentado a comer en una residencia cerrada, situada junto a una cafetería reservada para un programa para adolescentes. El conserje había seguido las normas, según las cuales había que avisar a seguridad si aparecía alguien que no perteneciera a ese programa. El guardia había mantenido una conversación educada con la estudiante y no le pidió que se fuera. Los trabajadores de la cafetería no habían tenido ninguna implicación en el asunto.

No obstante, cuando salió a la luz la verdad —en un extenso informe elaborado por un despacho de abogados externo que había contratado la universidad—, la versión de los hechos que había aparecido en Facebook, cimentada por la oleada de intensas emociones colectivas, ya hacía tiempo que se había asentado en las cabezas de la gente. La universidad, probablemente temerosa de generar más indignación, se negó a exonerar a los trabajadores declarando que era «imposible descartar la posible influencia de un sesgo racial implícito» en su conducta. A uno lo trasladaron, a otro se le mantuvo de baja y otro dejó el trabajo. Tiempo después, a una de las trabajadoras de la cafetería no la contrataron en un restaurante porque el entrevistador la reconoció de Facebook y la identificó como «la racista». La estudiante, primero alabada como una heroína, luego fue condenada como la mala de la película. A poca gente se le ocurrió culpar a las redes sociales de haber hecho posible que una adolescente destrozase la vida de unos trabajadores con ingresos bajos, de haberla incentivado a ella y a miles de usuarios a hacer tal cosa y de haber

permitido que todo el mundo percibiese aquella impresión falsa generada por la indignación como algo más verdadero que la verdad.

Que algo sea verdad o mentira tiene poca importancia al recibirse una publicación; la única diferencia es que un mentiroso tiene más libertad para alterar los hechos y adaptarlos a un relato que genere clics. Lo que importa es si la publicación puede provocar una reacción potente, por lo común de indignación. En un estudio hecho en 2013 sobre la plataforma china Weibo, se llegó a la conclusión de que, por regla general, la rabia viaja más lejos que otras emociones.²⁹ En estudios sobre Twitter y Facebook se ha llegado varias veces a esa misma conclusión, aunque los investigadores han restringido el efecto de la rabia en general a la indignación moral en concreto.³⁰ Los usuarios interiorizan las recompensas derivadas de la atención que proporcionan tales publicaciones, lo cual los lleva a producir más, lo que a su vez entrena los algoritmos de las plataformas para que fomenten aún más esas publicaciones.

Muchos de esos incidentes tenían un carácter de izquierdas, lo cual despertó temores de estar ante una «cultura de la cancelación» desbordada. Pero eso no hacía más que reflejar la concentración de usuarios de izquierdas en los espacios académicos, literarios, periodísticos y de otro tipo que tienden a ser más visibles en la sociedad estadounidense. También se estaba desplegando ese mismo patrón en comunidades de derechas. Pero a la mayor parte de esos casos no se les daba importancia por considerarlos obra de friquis (el Gamergate, los antivacunas) o de extremistas (los íncels, la derecha alternativa). Ya fueran de derechas o de izquierdas, el denominador común eran siempre las redes sociales, los incentivos que imponen y el comportamiento que provocan.

Para sus víctimas, los perjuicios, merecidos o no, son reales y duraderos. Nuestro cerebro procesa el ostracismo social literalmente como el dolor. Ser rechazado duele por la misma razón por la que duele un cuchillo que nos perfora la piel: la evolución nos ha hecho experimentar ambas cosas como una amenaza mortal. Nuestra sensibilidad social evolucionó mientras vivíamos en tribus, puesto que hacer enfadar a unos cuantos compañeros podía significar un riesgo real de morir. En las redes sociales, una persona puede —apenas sin previo aviso— sufrir la furia y la condena de miles de personas. A esa escala, los efectos pueden ser devastadores desde un punto de vista

psicológico. «Lo duro del acoso que las personas que no han sido acosadas en reiteradas ocasiones por una muchedumbre llena de odio tienen la suerte de no entender es que te cambia la vida para siempre —escribió una vez Pao, la exdirectora de Reddit—. Dejas de confiar con facilidad.»

Las consecuencias se extendían más allá de grupos de personas que eran víctimas de una rabia que se podría calificar de fuera de lugar o desproporcionada. La propia vida pública se estaba volviendo más ferozmente tribal, más extrema, más centrada en odiar y castigar cualquier mínima transgresión. «Te lo aseguro: estas plataformas no están diseñadas para mantener conversaciones razonadas —dijo Wu—. Twitter, Facebook y las redes sociales están diseñadas así: “Nosotros tenemos razón. Ellos están equivocados. Insultemos a esa persona lo antes posible y lo más duro que podamos”. Y no hacen más que amplificar todas las divisiones que tenemos.»

LOS ZORROS DE LIUDMILA

El misterio de la indignación moral —¿por qué nos atrae tanto una emoción que nos hace comportarnos de maneras que deploramos?— fue desvelado por una genetista rusa de setenta años escondida en un laboratorio de investigación en Siberia criando miles de zorros.³¹ Liudmila Trut llegó al laboratorio en 1959, recién graduada de la Universidad Estatal de Moscú, para investigar los orígenes de algo que hasta la fecha parecía no estar relacionado con ninguna otra cosa: la domesticación de los animales.

La domesticación era un misterio. Charles Darwin había especulado con que podía ser algo genético. Pero nadie sabía qué presiones externas convertían los lobos en perros o cómo la biología del lobo había cambiado para hacerlo tan amigable. Los discípulos de Darwin, no obstante, habían identificado una pista: todos los animales domesticados —ya fueran perros, caballos o vacas— tenían la cola o el rabo más cortos, las orejas más blandas, una constitución más delgada y un pelaje con más manchas que sus homólogos salvajes. Y muchos tenían una mancha característica en forma de estrella en la frente.

Si Trut lograba domesticar individuos en un entorno controlado, podría aislar las causas del proceso. Su laboratorio, situado junto a

una fábrica de pieles en Siberia, comenzó con cientos de zorros salvajes. Trut puntuó cada ejemplar evaluando su afabilidad con los humanos, crio solo al 10 % más afable y luego repitió el proceso con los descendientes de esa generación. A la décima generación, en efecto, nació un zorro con las orejas blandas. Otro tenía una marca en forma de estrella en la frente. Y, según escribió Trut, estaban «deseosos de establecer contacto humano, gimoteaban para atraer la atención y olisqueaban y lamían a los investigadores como los perros».³² Darwin tenía razón. La domesticación era algo genético. Las siguientes generaciones de zorros, a medida que se volvían más afables, tenían las patas, los rabos y los hocicos más cortos, el cráneo más pequeño, la cara más plana y más manchas en el pelaje.

Trut estudió a los animales durante medio siglo y terminó descubriendo el secreto de la domesticación: las células de la cresta neural. Cada animal inicia su vida con un conjunto de células de esta clase. Las células migran a través del embrión a medida que el animal crece y se convierten en los huesos de la mandíbula, en cartílagos, dientes, pigmento de la piel y partes del sistema nervioso. Por eso los zorros domesticados tenían marcas blancas en la frente: las células de la cresta neural que les habían legado sus progenitores, más afables, nunca llegaron ahí. Eso también daba cuenta de las orejas blandas, los rabos más cortos y los hocicos más pequeños.

Además, provocaba un cambio de personalidad, porque las células de la cresta neural también se convertían en las glándulas que producen las hormonas responsables de desatar el miedo y la agresión. Los zorros salvajes tenían miedo a los humanos y eran agresivos unos con otros, rasgos que les eran muy útiles en estado salvaje. Cuando Trut crio a los zorros más afables, sin saberlo estaba promoviendo a los animales con menos células de la cresta neural, con lo cual atrofió su desarrollo neurológico de una forma muy concreta y poderosa.

De las muchas revelaciones derivadas de las investigaciones de Trut, quizás la más destacada fue resolver un misterio de larga tradición sobre los humanos.³³ Hace unos 250.000 años, nuestro cerebro, tras haber aumentado durante millones de años, empezó a encogerse. Por extraño que parezca, eso ocurrió justo cuando los humanos parecían volverse más inteligentes, a juzgar por las herramientas encontradas con sus restos. Al mismo tiempo, los humanos desarrollaron unos huesos de brazos y piernas más delgados,

una cara más plana (se terminaron las frentes pronunciadas de los hombres primitivos) y unos dientes más pequeños, y también se volvieron más parecidos el cuerpo del macho y de la hembra. A la vista de los descubrimientos de Trut, de repente el motivo estaba claro. Esos eran los indicadores de una caída repentina de las células de la cresta neural: de la domesticación.

Sin embargo, los zorros de Trut habían sido domesticados por una fuerza externa: ella. ¿Qué había intervenido en la trayectoria evolutiva de los humanos para favorecer de la noche a la mañana a los individuos dóciles antes que a los agresivos? El antropólogo inglés Richard Wrangham encontró una respuesta: el lenguaje. Durante millones de años, nuestros antepasados que acabarían convirtiéndose en el *Homo sapiens* formaron pequeñas comunidades lideradas por un macho alfa. El macho más fuerte y más agresivo dominaba, y pasaba los genes en detrimento de los machos más débiles.

Todos los grandes simios detestan a los matones. Los chimpancés, por ejemplo, prodigan un trato preferencial a aquellos iguales que son amables con ellos y perjudican a aquellos que son crueles. Pero no tienen cómo compartir esa información entre ellos. Los matones nunca se resienten de su mala reputación porque, sin el lenguaje, no existe nada por el estilo. Eso cambió cuando nuestros antepasados desarrollaron un lenguaje lo bastante sofisticado para hablar del comportamiento de los demás. La agresividad dejó de ser un activo — el medio por el cual los machos alfa dominaban su clan— y pasó a ser una carga que el grupo, cansado de ser tratado con prepotencia, podía castigar si se unía.

«Las conspiraciones basadas en el lenguaje fueron la clave, porque eso daba a los machos beta susurradores el poder de unir fuerzas para matar a los machos alfa, los matones»,³⁴ escribió Wrangham en un revolucionario libro de 2019. Cada vez que un clan de humanos primitivos derrocaba a un alfa despótico, hacían lo mismo que hizo Liudmila Trut a sus zorros: seleccionarlos según su docilidad. Los machos más cooperativos se reproducían; los agresivos no. Nos autodomesticamos.

No obstante, al mismo tiempo que los primeros humanos ahuyentaban una forma de agresividad, estaban seleccionando otra: la violencia colectiva que habían empleado tanto para derribar a los alfas como para imponer un nuevo orden en su lugar. La vida empezó

a estar gobernada por lo que el antropólogo Ernest Gellner llamó la «tiranía de los primos».³⁵ Las tribus dejaron de tener líderes, se convirtieron en sociedades basadas en el consenso, unidas por la fidelidad a un código moral compartido, que los adultos del grupo (los «primos») hacían cumplir, a veces con violencia. «No actuar conforme a las normas, contravenir los estándares de la comunidad o granjearse una reputación de mala persona se convirtió en una aventura peligrosa»,³⁶ escribió Wrangham. Si ofendes al colectivo, puedes acabar ignorado o exiliado, o con una pedrada en la frente. La mayoría de las sociedades de cazadores-recolectores viven de esta forma hoy en día, lo cual hace pensar que la práctica se basa en algo intrínseco de nuestra especie.

La base de ese nuevo orden era la indignación moral. Así alertabas a tu comunidad de un mal comportamiento. Así la convocabas —o se te convocaba— para ir a castigar una transgresión. Y era la amenaza que sobrevolaba tu cabeza desde que nacías hasta que morías, y que te mantenía a raya. La indignación moral, cuando cobra suficiente fuerza, se convierte en lo que Wrangham llama agresión «proactiva» o «coaligada»,³⁷ lo que en el lenguaje coloquial se conoce como una turba. Cuando ves a una turba, en esa muchedumbre violenta estás viendo la tiranía de los primos, el mecanismo de nuestra autodomesticación. Esa amenaza, a menudo letal, se convirtió en una presión evolutiva por sí sola, lo que nos llevó a desarrollar una sensibilidad aguzadísima ante los estándares morales del grupo. Y el instinto de ceñirnos a ellos. Si quieres demostrarle al grupo que puede confiar en ti para hacer cumplir sus estándares, agarra una piedra y arrójala. Si no, tú podrías ser el siguiente.

En nuestra historia más reciente, decidimos que esos impulsos son más peligrosos que beneficiosos. Sustituimos la tiranía de los primos por el imperio de la ley (en la mayoría de los casos), prohibimos la violencia colectiva y desalentamos que la gente se comportase como una muchedumbre enfurecida. Pero los instintos no pueden neutralizarse por completo, solo pueden contenerse. Las redes sociales, sacando partido directamente de nuestras emociones de grupo más viscerales, esquivan ese muro de contención y, dadas las circunstancias adecuadas, lo derriban del todo, con lo cual esos comportamientos instintivos se desparan de nuevo por la sociedad.

Cuando ves una publicación que expresa indignación moral,

entran en acción 250.000 años de evolución. Te sientes empujado a sumarte al grupo. Eso hace que olvides tus principios morales internos y te unas a los del grupo. Y hace que infligir daño a la persona que es objeto de la indignación parezca algo necesario, incluso de lo más placentero. En escáneres cerebrales se ha visto que, cuando los sujetos hacen daño a alguien que creen que es un infractor moral, se activan sus centros de recompensa de la dopamina.³⁸ Las plataformas también eliminan muchos de los controles que por lo general nos frenan para impedir que las cosas vayan demasiado lejos. Desde detrás de una pantalla, lejos de nuestras víctimas, no sentimos ni un ápice de culpa al ver el dolor en el rostro de la persona a la que hemos hecho daño. Ni tampoco sentimos vergüenza al darnos cuenta de que nuestra rabia ha degenerado visiblemente en crueldad. En el mundo real, si gritas improperios a alguien por llevar una gorra de béisbol en un restaurante caro, el que va a ser despreciado vas a ser tú, castigado por violar las normas que prohíben las muestras de rabia excesivas y por molestar a los demás clientes del restaurante. En internet, por el contrario, si otras personas se dan cuenta de tu arrebató, probablemente acabarán uniéndose a ti.

Las plataformas sociales cuentan con una abundancia antinatural de fuentes de indignación moral: siempre hay un tuit o una noticia para enfadarse, y otros muchos usuarios con quienes destacar ese elemento a un público potencial de millones de personas. Es como estar en el centro de la mayor multitud reunida jamás sabiendo que en cualquier momento podría transformarse en una muchedumbre violenta. Eso crea fuertes incentivos para lo que los filósofos Justin Tosi y Brandon Warmke han denominado «fanfarronería moral»: presumir delante de los demás de que estás más indignado y, por tanto, eres moralmente superior al resto. «Queriendo impresionar a tus iguales —escriben Tosi y Warmke—, los fanfarrones se inventan acusaciones morales, se suman a casos ya iniciados de humillación pública, proclaman que cualquier persona que no esté de acuerdo con ellos está equivocada de forma evidente o exageran sus manifestaciones emocionales.»³⁹

Los fanfarrones morales en el mundo físico podrían agudizar un poco las sensibilidades de un grupo concreto presionando a sus iguales para que les siguieran la corriente. O sencillamente podrían sacar de quicio a todo el mundo. Pero, en las redes sociales, los fanfarrones se

ven recompensados y amplificadas de un modo sistemático. Eso puede desencadenar «una carrera armamentística moral», alertaron Tosi y Warmke, en la que las personas «adopten visiones extremas e inadmisibles y se nieguen a escuchar a la parte contraria». ⁴⁰

Si eso no ocurriese más que en algunos foros de internet, las consecuencias serían algunas discusiones desagradables. Pero, a mediados de la década de 2010, las redes sociales se habían convertido en el vector a través del cual se consumían e interpretaban buena parte de las noticias del mundo. Eso creó un mundo — advirtieron Tosi y Warmke en un estudio de seguimiento con el psicólogo Joshua Grubbs— definido por «la homogeneidad, los sesgos de endogrupo/exogrupo y una cultura que promueve la indignación». ⁴¹

El resultado fue un bucle fatídico de polarización y desinformación. En 2020, cuando el Congreso de Estados Unidos aprobó un paquete de estímulo, las publicaciones más compartidas en Twitter informaron de que el proyecto de ley desviaba 500 millones de dólares que de entrada estaban destinados a estadounidenses con bajos ingresos al Gobierno de Israel, y se apropiaba de otros 154 millones que tenían que ir destinados a la Galería Nacional de Arte; que financiaba una operación clandestina de 33 millones de dólares para derrocar al presidente de Venezuela; que recortaba las prestaciones por desempleo, y que, en realidad, los cheques de 600 dólares de ayuda para el covid no eran más que unos préstamos que la agencia tributaria estadounidense (IRS) recuperaría mediante los impuestos del año siguiente.

Todas esas informaciones eran falsas. Pero la inclinación extrema de la plataforma hacia la indignación hizo que la desinformación se impusiera, lo cual creó una demanda de más rumores y mentiras en que se afirmase la indignación. Los insensibles republicanos querían que los pobres murieran de hambre. Los cobardes demócratas habían vendido a los estadounidenses a las grandes empresas. Unos astutos extranjeros habían robado nuestra savia financiera. Cada ciclo dificultaba aún más que los ciudadanos comprendieran un asunto capital y hacía más costoso y menos factible que los legisladores llegasen a acuerdos, cuyos votantes les exigían nada menos que posiciones partidistas maximalistas para calmar su rabia.

Esos niveles de rabia crean un impulso, a veces acuciante, de

encontrar a alguien a quien castigar. En un experimento perturbador, se pidió a los participantes que asignasen un castigo para la transgresión moral de otra persona.⁴² Tomaban decisiones más duras cuando les hacían creer que los estaban observando, y aún más duras cuando se les decía que su público estaba muy politizado o ideologizado. Muchos endurecían el castigo incluso si pensaban que su víctima no lo merecía. Su motivación era simple: tenían la expectativa de que la crueldad haría que los observadores los quisieran más.

Y ese efecto aumenta con la escala: las personas expresan más indignación y demuestran una mayor voluntad de castigar a quienes no lo merecen cuando piensan que su público es mayor. Y no hay mayor público sobre la faz de la Tierra que Twitter o Facebook.

UN PASEO POR THE RAMBLE

Si había alguna duda acerca del alcance de la influencia de las humillaciones por Twitter sobre la sociedad estadounidense, esas dudas se disiparon el Día de los Caídos de 2020 en una extensa y tranquila zona verde del Central Park de Nueva York conocida como The Ramble.⁴³ Casi todas las mañanas, en esa área se repite un mismo ritual. Alguien deja correr libremente a su perro, un respiro momentáneo para dejar atrás los pisos estrechos y las ajetreadas aceras. Y alguien que ha venido a observar pájaros, que considera que esa es una zona magnífica, reprende al propietario del perro para que vuelva a poner la correa al animal, como exigen las normas del parque.

Esa mañana, el propietario del perro era una mujer blanca, Amy Cooper. El amante de los pájaros era un hombre negro, Christian Cooper, quien había vivido ese ritual tantas veces que llevaba comida para perros para atraer a los animales sueltos, lo cual forzaría a los dueños a atarlos de nuevo. Cuando el perro de Amy se acercó a la zona de Christian, él pidió a la mujer que lo atase. Ella se negó.

Él hizo un gesto enseñando la comida para perros con la que atraería al animal y le dijo a la mujer:

—Mire, si usted hace lo que le da la gana, yo también voy a hacer lo que me da la gana, pero no le va a gustar.

—Ni se le ocurra tocar a mi perro —gritó ella.

Intuyendo que el conflicto subiría de tono, el hombre sacó el móvil, apuntó la cámara hacia la mujer y empezó a grabar. Ella avanzó hacia él y le pidió que dejase de filmar, mientras él le decía que no avanzase más, y de este modo los dos intentaban dominar la situación. Dos urbanitas de mediana edad, de clase media alta, con la voz temblándoles por el inhabitual subidón de adrenalina.

—Pues entonces le haré una foto y llamaré a la policía —dijo ella.

—Sí, llame a la policía —dijo él—. Adelante, llámela.

—Les voy a decir que hay un hombre afroamericano poniendo mi vida en riesgo —dijo ella.

Aquello era una amenaza potencialmente mortal. Durante buena parte de esa primavera habían aparecido noticias de hombres y mujeres negros muertos a manos de la policía. Amy llamó al 911 y le dijo al operador en reiteradas ocasiones que «un hombre afroamericano» la estaba amenazando. «Por favor, envíe a la policía de inmediato», dijo, levantando el tono de voz en lo que pareció un intento de fingir terror. A media llamada, volvió a atar a su perro, y Christian, al recobrar la calma, le dio las gracias secamente y dejó de grabar.

En un mundo sin redes sociales, el incidente tal vez habría terminado ahí. Cuando llegó la policía, ambas personas se habían ido del parque. Christian podría haber utilizado el vídeo para acusar a la mujer de haber falseado un aviso a la policía. Pero la mayor transgresión de Amy —blandir una amenaza de violencia policial para amedrentar a Christian y poner en peligro su vida— habría quedado impune. Así como su intento de imponer la expectativa de que los espacios públicos como el parque pertenecen a los blancos antes que a los negros, punible bajo pena de muerte. En este mundo, sin embargo, la hermana de Christian pudo colgar el vídeo en Twitter. «Quería que la gente supiera lo que ha sucedido para asegurarme de que esa señora nunca vuelva a repetir ese comportamiento», dijo a sus pocos seguidores.

Justo para disuadir ese tipo de transgresión existe la indignación moral: una violación de las costumbres compartidas (no mientas, no pongas en peligro a los demás, no fomentes el racismo) y un ataque al contrato social que garantiza nuestra cohesión social. También demostraba una forma muy perniciosa de racismo, sobre la cual los

usuarios podían llamar la atención compartiendo el vídeo. Uno republicó el vídeo, luego otro y luego otro, hasta que más de 200.000 personas hubieron compartido la publicación, cada una de ellas declarando que estaban de acuerdo con que se exigiesen responsabilidades a Amy Cooper. La publicación atrajo 40 millones de visualizaciones, veinte veces la audiencia del telediario vespertino.

Decenas de miles de usuarios, hablando con una sola voz furiosa, presionaron a los jefes de Amy Cooper, quienes la despidieron de inmediato, y divulgaron los nombres y cuentas en las redes sociales de sus amigos, un llamamiento implícito a que la aislasen socialmente. Los ataques a Cooper llegaron incluso a la perrera donde había adoptado a su mascota, lo cual la llevó a entregar a su querido perro (la perrera se lo devolvió al cabo de una semana). Se hicieron públicos ante millones de personas detalles sobre su vida sórdidos y no vinculados con el caso. La habían traído a la plaza más grande del mundo de la historia de la humanidad y la habían condenado por unanimidad. Eso era hacer justicia, tanto para Christian Cooper como para dejar claro que una conducta como esa sería castigada de inmediato. E hizo que se dedicara una atención muy necesaria al poder que la violencia policial da a cualquier persona blanca con un móvil en la mano para amenazar a una persona negra.

Sin embargo, Twitter también llevó esa justicia justo hasta el extremo que los usuarios más furiosos del sitio quisieron llevar su rabia espoleada por el algoritmo. Aunque el colectivo virtual había alcanzado el veredicto correcto, algunos expresaron incomodidad con los métodos con los que había determinado y ejecutado la sentencia. Entre quienes expresaron cierto grado de ambivalencia estaba el propio Christian Cooper, quien, aunque no dejó de criticar a Amy Cooper por intentar dirigir «ciertos impulsos sociales oscuros» contra él, añadió: «No excuso el racismo, pero no sé si hacía falta destrozarle la vida».⁴⁴

El hecho de que se la presionara para que devolviera a su perro en el momento álgido de la furia virtual pero que lo recuperara cuando las aguas volvieron a su cauce, hacía pensar que las dinámicas de las redes sociales, al menos por un tiempo, llevaron su castigo más allá incluso de lo que los castigadores creían apropiado. Cualquier sistema tiene sesgos, puntos ciegos y excesos. Los Cooper demostraron que todos habíamos entrado en un nuevo sistema que había

suplantado a los viejos, sin que nadie lo hubiera diseñado a conciencia, nadie hubiese escogido acogerse a él o sin que nadie lo comprendiera del todo.⁴⁵

Hay veces en que ese sistema es transformador en un sentido positivo. Los activistas del movimiento Black Lives Matter lo aprovecharon para que la sociedad se fijara en una violencia que los medios mayoritarios tendían a tratar por encima. El vídeo de Christian Cooper tuvo una reacción tan potente en parte porque esos activistas habían preparado a millones de personas para que vieran su importancia. Ese mismo día, un policía de Mineápolis mató a un hombre negro llamado George Floyd poniéndole la rodilla encima del cuello durante casi nueve minutos. A lo largo de varias semanas, millones de personas participaron en manifestaciones en muchas ciudades que fueron la culminación de una organización sobre el terreno y también una fuente a nivel nacional de indignación moral que en buena medida había tenido lugar a través de las redes sociales. Las acusaciones de acoso sexual contra el productor cinematográfico Harvey Weinstein también habían generado un ciclo de indignación creciente en internet —primero contra él, luego contra Hollywood y luego contra los hombres agresores nombrados en innumerables historias personales compartidas en Twitter— que se convirtió en el movimiento MeToo. La promesa de revolución de Silicon Valley había sido algo interesado, medio planeado y solo un componente de una destabilización más general que a menudo era dañina, pero en la que había parte de verdad.

No obstante, los incidentes injustos también se han vuelto frecuentes. Fallos, errores, indignación expresada con finalidades perversas o sin motivo alguno. Como también los casos que caen en áreas morales grises: más dentistas de cincuenta y cinco años que tienen que esconderse o responsables de *marketing* de treinta y pico desposeídos de su fuente de ingresos. El cambio no se produjo porque de repente la sociedad se volviese más justa o por el ascenso de la llamada cultura de la cancelación; fue por la llegada de una tecnología tan omnipresente y tan arraigada en nuestra cognición que había alterado la manera de funcionar de la moralidad y la justicia. Los Coopers y los Walters Palmer y los Gamergates forman parte todos de ese mismo nuevo orden, nuestra tiranía de los primos digital.

Ese sistema automatizado y sensorialmente abrumador puede

explotarse con facilidad. Como había alertado James Williams, el exingeniero de Google, Trump ganó popularidad gracias a tuits ultravirales y publicaciones de Facebook fomentando la rabia contra demócratas, periodistas y minorías, a menudo por pecados inventados pero que inspiraban llamamientos muy reales a la detención o al asesinato de sus objetivos. Dedicad una hora a navegar por un sector de la web social que no comparta vuestras ideas políticas y preguntaos si la indignación omnipresente es extrapolable y si realmente los castigos que se reclaman siempre son apropiados.

Algunas semanas después del enfrentamiento en The Ramble, un hombre de Seattle llamado Karlos Dillard colgó en Twitter un vídeo de dos minutos que presentaba como una secuela. Decía que una «Karen» (una forma coloquial de referirse a una mujer blanca que se cree con privilegios, un término que también había usado la hermana de Christian Cooper) le había cerrado el paso en la carretera mientras gritaba insultos racistas. Incluso el encuadre era el mismo que el del vídeo de The Ramble, aunque en este caso tenía lugar en la entrada de domicilio de la mujer, donde Dillard, un hombre negro, dijo que la había seguido para pedirle explicaciones.

—No entiendo qué está ocurriendo —decía la mujer chillando en los primeros segundos del vídeo, con las manos temblando mientras intentaba taparse la cara.

—Me cierra el paso y ahora se hace la víctima —decía Dillard mientras hacía una panorámica del bloque de pisos de la mujer—. Chicos, aquí podéis ver su número de matrícula. Vive aquí. Esta es su dirección.

Ella se echó encima de la matrícula y le suplicó que dejase de grabar. Dillard se detuvo como si estuviera pensando y, gritando para que lo oyera todo el mundo, le exigió que se disculpase por referirse a él con la palabra *nigger* ('negro'), un término muy ofensivo en Estados Unidos. Ella gritó, apenas pudiendo articular palabra:

—Vas a arruinarme la vida y ni siquiera me conoces.

Mientras sollozaba en la acera, él la acribilló con exigencias de que le pidiera perdón. Mientras la mujer se estaba negando, balbuceó:

—Si mi marido es negro...

El vídeo de Dillard se hizo viral al instante, compartido por más de 100.000 personas y con más de 10 millones de visualizaciones. Los usuarios denunciaban a la mujer, exigían que la echasen del trabajo y

difundían información para identificarla. Muchos se regodeaban. «LMAOOOOO, qué dramática, la tía —escribió uno, usando un acrónimo de internet que indica una risa intensa—. Que quede retratada.» Otros encontraron a su marido y difundieron su nombre e imagen como represalia por lo que decían que era el problema de los hombres negros que hacían posible que hubiese mujeres blancas racistas. Muchos de los comentarios más furibundos procedían de usuarios blancos: quizás sinceros, quizás fanfarrones, quizás ambas cosas.

Dillard aprovechó su fama y creó una página web en la que vendía camisetas y sudaderas con citas del vídeo. Tiempo después, alguien descubrió que anteriormente había colgado vídeos en los que se enfrentaba airado con mujeres acusándolas de racismo, con lo cual generaba negaciones confusas y temerosas. Como mínimo una había obtenido una orden de alejamiento. En otro vídeo antiguo, Dillard había presumido de inventarse acusaciones de racismo. «Eso fue una mentira que inventé en ese momento —dijo—. Se llama “Tiende una trampa”.» En otro, decía que había votado a Trump. Varios periodistas entrevistaron a Dillard, quien resultó ser un personaje confuso e incoherente, no un experto manipulador ni un genio cínico. Simplemente había aprendido, después de repetirlo unas cuantas veces, qué combinación de palabras e imágenes le reportarían validación y atención en la mayor máquina de generar indignación de la historia. Las plataformas se ocupaban del resto.

Eso fue solo un paréntesis en medio de un verano en que, por otro lado, los activistas serios dirigieron la maquinaria de la indignación virtual hacia fines más dignos. Pero no era su maquinaria. Aunque pudiera parecer gobernada por la voluntad colectiva de sus participantes, en realidad estaba gobernada por Silicon Valley, cuyos sistemas no se diseñaron para promover el progreso social o para distribuir la justicia de forma equitativa, sino para maximizar el tiempo pasado en las plataformas, para ganar dinero.

En política, los resultados pocas veces privilegiaban la liberación. Cuando dos expertos analizaron 300 millones de tuits publicados durante la campaña presidencial de 2012, descubrieron que los tuits falsos habían superado a los verdaderos de forma sistemática.⁴⁶ Los rumores y las mentiras —alertaban los expertos— fomentaban la rabia del otro bando, lo cual ensanchaba la polarización, que ya era uno de

los males más graves a los que se enfrentaba la democracia estadounidense. La división resultante estaba allanando el camino a los oportunistas. Los expertos se dieron cuenta de que uno de los peores motores de la desinformación en Twitter durante las elecciones fue un personaje televisivo, entonces marginal, llamado Donald Trump. Aun así, la provocación viral, sin importar lo mucho que proliferase, ejerció, por su cuenta, una gran influencia. Trump había dominado Twitter, pero poco más. Si las plataformas hubieran permanecido estáticas, esas oleadas de indignación y conflicto, a pesar de su fuerza distorsionadora y algunas veces destructiva, podrían haber supuesto el punto álgido del impacto de las redes sociales. Pero un conjunto de avances tecnológicos reforzaría el poder de las plataformas hasta tales extremos y a un ritmo tan rápido que, en las siguientes elecciones, el propio mundo se reconfiguraría a su imagen y semejanza.

Despertar a la máquina

LA ALGORITMIFICACIÓN

Hace mucho tiempo —tanto que podría haberse evitado buena parte del caos que estaba por venir—, un especialista en inteligencia artificial (IA) de Google intentó echar un vistazo a uno de los mayores secretos a voces de su sector: nadie sabe del todo cómo funcionan realmente los algoritmos que rigen las redes sociales. Estos sistemas operan de forma semiautónoma, con unos métodos que los humanos no llegan a entender. Pero Silicon Valley tenía un incentivo para permanecer en la ignorancia. Si vas a mirar cómo la gallina consigue los huevos de oro, puede que no te guste lo que encuentres. Tal vez incluso podrías tener que devolverlos.

El especialista en cuestión, un francés llamado Guillaume Chaslot, admiraba a las empresas de redes sociales desde que era estudiante de doctorado en Europa en la década del año 2000. La tecnología de las plataformas no era compleja, pero su potencial, creía Chaslot, era revolucionario. Una vez terminada su tesis, se marchó a California. A finales de 2010, consiguió un empleo en Google.

«No sabía en qué iba a trabajar, porque ellos simplemente contratan personas y luego las ponen en un proyecto», me contó Chaslot, que habla con un susurro enérgico. Al final le dijeron que trabajaría en una plataforma de vídeo que Google había adquirido a petición de una ejecutiva publicitaria llamada Susan Wojcicki. En 1998, Larry Page y Sergey Brin, los creadores y cofundadores de Google, habían montado los primeros servidores de la empresa en el garaje de Wojcicki. Cuando su motor de búsqueda se puso de moda, Wojcicki (pronunciado «woujiski») dejó su trabajo en Intel para trabajar para ellos. Supervisaba productos publicitarios y un servicio de *streaming*, Google Videos, que estaba siendo superado en una proporción de tres a uno por una empresa emergente muy simple llamada YouTube. Wojcicki, que creía que ya era demasiado tarde

para seguirle el ritmo, convenció a sus jefes de comprarla, lo cual hicieron, en 2006, por la asombrosa cifra de 1.650 millones de dólares.

Pese a sus proyecciones de crecimiento, los ingresos publicitarios nunca aumentaron lo bastante para justificar el coste. Cuatro años después, con la esperanza de salvar la inversión, Google contrató a varios programadores muy especializados, entre ellos Chaslot. Encargaron la dirección del proyecto a un veterano de los motores de búsqueda llamado Cristos Goodrow. «En septiembre de 2011, envié un correo provocador a mi jefe y al equipo de dirección de YouTube — escribió más adelante Goodrow—. Asunto: “Tiempo de visualización y nada más que tiempo de visualización”. Era un llamamiento a repensar cómo evaluábamos el éxito.»¹

Goodrow pidió a sus jefes que imaginaran a un usuario hipotético que buscase cómo ponerse una pajarita. Imaginad que YouTube — decía en su mensaje— pudiera mostrarle al usuario un vídeo enseñándole cómo hacer eso en un minutito. O que pudiera mostrarle un vídeo que «dure diez minutos y que esté lleno de chistes o que sea muy entretenido, y que al final te hayas enterado —o no— de cómo se pone uno una pajarita».² La ortodoxia de Google optaba por mostrar el primer vídeo: tiene que aparecer en primer lugar la información más útil lo más deprisa posible. Pero Goodrow sostenía que YouTube tenía que fomentar el segundo. «Nuestra misión era hacer que los usuarios participasen y se quedasen en nuestra plataforma», escribió. Dar a los usuarios un vídeo largo que no quieran quitar, luego otro, y otro más. Un mayor tiempo de visualización «genera más publicidad, lo cual incentiva que haya más creadores de contenido, lo cual genera más público»,³ argumentó. Sus jefes estuvieron de acuerdo.

Chaslot se encargó de un componente esencial de ese proyecto: la búsqueda. Las búsquedas tradicionales se basan en palabras clave: si tecleas «ballenas», te sale una lista de los vídeos más recientes o más vistos etiquetados con esa palabra. El equipo de Chaslot sustituiría eso con una inteligencia artificial diseñada para identificar el vídeo que se adecuara mejor a los intereses del usuario. Buscar «ballenas», en teoría, haría que la inteligencia artificial rastreara los miles de millones de horas de vídeo de YouTube para encontrar un documental sobre Jacques Cousteau que parezca una joya escondida o un vídeo *amateur* impresionante de una orca saliendo a la superficie. Incluso

podía recomendar qué ver a continuación, con lo cual guiaría a los usuarios a través de un mundo infinito de descubrimientos y placer. «Fue una tarea —me dijo Chaslot— que tuvo unos enormes efectos positivos en la vida cotidiana real de muchísima gente.»

Chaslot vio por qué le habían encargado esa labor. La nueva búsqueda tendría que hacer el trabajo de un ejecutivo de televisión de primera categoría, medir los gustos y preferencias de las audiencias, pero seleccionar de entre un conjunto de vídeos millones de veces mayor que el de cualquier cadena de televisión, y todo ello a una velocidad casi instantánea. Chaslot conocía —gracias a sus investigaciones doctorales, que versaban sobre algo llamado aprendizaje automático— una tecnología que poco antes había resuelto un problema irresoluble hasta el momento: el *spam* o correo basura.

Los primeros filtros de *spam* podían identificar los correos basura solo basándose en identificadores a los que se les hubiera dado instrucciones de buscar, por ejemplo una dirección de correo que ya se sabía que era un remitente basura o ciertas palabras clave. Pero los supervisores de los filtros de *spam* tenían que identificar y programar esos indicadores personalmente. Los correos masivos automatizados podían colapsar esas defensas, lo cual daba riqueza fácil a los remitentes basura. En 2002, el correo basura representaba el 40 % de todos los emails e iba en aumento.⁴ Las guerras del *spam* parecían imposibles de ganar. Los supervisores de internet avisaron de que el tráfico de *spam* pronto aumentaría hasta el punto de dejar el correo electrónico inservible y tal vez provocaría la quiebra de internet.

Con el aprendizaje automático, los ingenieros podían hacer algo mejor que escribir un programa para identificar correos basura. Diseñaban un programa que guiaría su propia evolución. Alimentaban ese programa con enormes cantidades de correos basura y no basura. Luego el sistema creaba de forma automática miles de filtros de *spam*, todos un poco diferentes, y los testeaba con emails de muestra. Entonces creaba una nueva generación de filtros de *spam* basados en los que habían dado mejor resultado y repetía el proceso, una y otra vez, como un botánico identificando y cruzando las plantas más resistentes. Iba evolucionando a una velocidad vertiginosa hasta que producía una variación de sí mismo tan sofisticada y poderosa que podía hacer lo que ningún otro filtro diseñado por los humanos era

capaz de hacer: identificar y bloquear proactivamente casi todo el correo basura. Un supervisor no tiene forma de abrir el capó de ese filtro de *spam* y ver cómo funciona, porque estaría observando una máquina que, con el tiempo, fue diseñada por otra máquina, y que es demasiado compleja para que la comprendamos. Pero ¿a quién le importa? Esas máquinas derrotaban con facilidad al correo basura, lo cual salvó internet de la catástrofe.

Google, Facebook y otras empresas absorbieron a las mejores figuras del ámbito del aprendizaje automático. A muchas se les asignó una labor parecida a la de Chaslot. En lugar de identificar *spam*, debían crear máquinas que aprendiesen con precisión qué combinaciones de texto, imágenes y sonidos serían mejores para hacer que los usuarios siguiésemos desplazándonos pantalla abajo.

Puestos en funcionamiento a principios de 2012 en YouTube, los poderes de ese nuevo sistema se extendieron más allá de los meros resultados de búsqueda. Imaginad que veis, por ejemplo, un vídeo de un debate presidencial de 2012. La página ahora os recomendaría, al lado del vídeo, las miniaturas de una docena de vídeos para ver a continuación: un vídeo de las peores pifias de Obama, una parodia de *Saturday Night Live*, un bloguero denunciando las políticas de Mitt Romney. Una vez terminado el vídeo que estéis viendo, el sistema incluso escogerá uno de esos vídeos para que se reproduzca de forma automática justo después. Cada vídeo es seleccionado de entre los miles de millones de vídeos de YouTube por una inteligencia artificial corporativa con el nombre en clave de «el algoritmo»: uno de los sistemas de aprendizaje automático más potentes del sector de la tecnología de consumo. Sus selecciones, guiadas por el poder del aprendizaje automático, resultaron tener una enorme eficacia. «En pocos meses, con un pequeño equipo, disponíamos de un algoritmo que aumentaba el tiempo de visualización para generar millones de dólares de ingresos publicitarios más —dijo Chaslot—, así que era muy, pero que muy emocionante.»

Esos sistemas estaban adentrándose con sigilo en todos los ámbitos de la vida. El de Netflix, por ejemplo, aprende los gustos de cada usuario rastreando qué ve la persona y cuánto tiempo, y luego hace recomendaciones tan efectivas que la empresa responsabiliza a su algoritmo de una fidelización de suscriptores que les reporta mil millones de dólares cada año.⁵ Spotify adquirió empresas de IA para

crear algoritmos que seleccionasen listas de reproducción que generan buena parte de su negocio, de 8.000 millones de dólares anuales.⁶ Si compras en Amazon, un algoritmo busca datos de hábitos de compra para determinar qué productos te va a mostrar. Si lees Google News, un algoritmo decide qué titulares te van a apetecer más. Incluso el amor ha acabado en manos de los algoritmos de las aplicaciones de citas, las cuales reducen el encanto y las esperanzas de los usuarios a meros datos, que ese programa utiliza para alentar a las personas a aparearse.

El sistema de YouTube busca algo con un alcance mucho mayor que una cuota mensual de subscripción. Su ojo panóptico rastrea cada detalle de lo que ves, cuánto rato lo miras y qué clicas a continuación. Monitoriza a 2.000 millones de usuarios, con lo cual acumula lo que es, sin lugar a dudas, el conjunto de datos más grande sobre preferencias de visualización jamás recopilado, y lo examina constantemente para encontrar patrones. Chaslot y otros empleados iban modificando el sistema sobre la marcha, alentando su proceso de aprendizaje para alcanzar mejor su objetivo: maximizar el tiempo de visualización.

Una de las herramientas más potentes del algoritmo es la afinidad temática. Si ves un vídeo de gatos hasta el final, explicó Chaslot, YouTube te recomendará más cuando vuelvas a la plataforma. Dará especial preponderancia a aquellos vídeos de gatos que haya considerado más eficaces a la hora de captar la atención. Por ejemplo, una larga compilación de escandalosas meteduras de pata de gatitos. Como casi todos los usuarios de internet, yo también he experimentado eso. Los fines de semana voy en bici y, mientras vivía en el extranjero, buscaba vídeos de YouTube para ver rutas por la zona y hacerme una idea del terreno. El sistema empezó a recomendarme vídeos de ciclismo que nunca habría pensado que buscaría: carreras profesionales, carreras de prueba de modelos nuevos. Funcionó; vi más vídeos. Con el tiempo, las recomendaciones se volvieron más extremas. Choques dramáticos, colisiones múltiples de diez bicis, escenas de riesgo desafiando a la muerte. Aunque no eran nocivas ni mucho menos, tampoco eran muy divertidas; pero eran inusualmente cautivadoras, como un accidente de tráfico. Y, en algunos casos, eran eso mismo. El efecto es empujar a los usuarios hacia variaciones cada vez más emocionantes de sus intereses. Si son

gatos o bicis, las consecuencias son mínimas. Si se trata de asuntos políticos, de salud u otras cuestiones con cierta relevancia para la sociedad, las consecuencias pueden ser notables.

A medida que el sistema afinaba sus capacidades, Chaslot se dio cuenta de que desarrollaba unos extraños hábitos. Empezó a alentar a muchos usuarios a ver vídeos que fomentaban la rabia contra las mujeres. A veces contra mujeres concretas, como Anita Sarkeesian, crítica con la cultura de los videojuegos. A veces contra las mujeres en general. Los hombres pasaban un 40 % más de tiempo que las mujeres en YouTube, herencia en parte de la enorme cantidad de contenido relacionado con los videojuegos que había en la plataforma en aquel entonces.⁷ Chaslot se dio cuenta de que lo natural era que el algoritmo privilegiase un contenido más centrado en los hombres.

Al igual que con Twitter y Reddit, la indignación y el tribalismo activan las emociones de los usuarios con mayor eficacia en YouTube, y eso los lleva a ver más y más vídeos: justo lo que Goodrow había pedido al equipo de Chaslot que priorizase. El algoritmo aprendió a incrementar el tiempo de visualización de los fans de los videojuegos mostrándoles un vídeo en el que se expresaba rabia antifeminista, luego otro y luego otro más. Los vídeos a menudo hacían referencia a hombres que no estaban seguros de qué opinión tenían acerca de las cuestiones de género, quizás por ningún otro motivo que su juventud. «Es un círculo vicioso —dijo Chaslot—. El problema de las madrigueras.» Aunque muchos usuarios no prestaban atención a los vídeos, suficientes personas se enganchaban a ellos para entrenar el sistema a ofrecer vídeos parecidos una y otra vez. Los creadores de los vídeos se percataron de que títulos como «LA VERDAD SOBRE EL FEMINISMO» generaban montones de visualizaciones, así que hicieron más.

Una de las libertades más aclamadas de Google, heredada de los fundadores de mediados de siglo de Silicon Valley y tomada prestada directamente de los programas de investigación de Stanford del otro lado de la ciudad, es la regla del 80/20. Los empleados dedican un 80 % de su tiempo a tareas formales, pero pueden dedicar el 20 % restante a llevar a cabo proyectos secundarios. Chaslot y el director de su equipo, que compartía sus preocupaciones, aprovecharon su 20 % para desarrollar un nuevo algoritmo que pudiera equilibrar los objetivos de beneficios con el bienestar de los usuarios.

Ese otoño, en 2012, en un congreso de directores de YouTube en

Los Ángeles, un ejecutivo llamó aparte a Goodrow y a unos cuantos empleados más para decirles que iba a hacer un anuncio sorpresa.⁸ La compañía se reorientaría en torno a un objetivo total: incrementar el tiempo de visualización diario por diez. Sus servidores ya registraban cien millones de horas de tiempo de visualización diario. Pero, aunque YouTube ya se estaba expandiendo a nuevos países y los espectadores televisivos iban pasándose de un modo paulatino al consumo por internet, las horas de visualización podían crecer igual de rápido. Había que seducir a los usuarios que tenían la intención de ver un vídeo para que se quedasen a ver muchos más. La capacidad de persuasión del algoritmo tenía que aumentar de forma drástica.

¿Cuándo podía lograrse eso?, le preguntó el ejecutivo. ¿Cuál era el marco temporal? Goodrow contestó que 2015 sería demasiado pronto. Pero 2017, escribió, «sonaba raro» porque era un número primo. Acordaron que lo lograrían a finales de 2016, cuatro años después, y más tarde Goodrow prometió dimitir si fracasaba. Aquello puso a YouTube en un esprint hacia un plazo autoimpuesto, en el que sus ejecutivos e ingenieros estarían dedicados por completo a privilegiar contenido que enganchara a los usuarios el mayor tiempo posible; y todo eso en paralelo a unas elecciones presidenciales en las que su influencia resultaría fatídica.

BURBUJAS DE OPINIÓN

Chaslot no era la única persona de Silicon Valley preocupada por las consecuencias de los algoritmos. Para referirse a esas preocupaciones, se había extendido una expresión desde el verano anterior, en 2011. Una mañana de mayo, mientras Chaslot trabajaba en su despacho en las oficinas de Google en Los Ángeles, sus jefes entraban en un salón de convenciones al otro lado de la ciudad, donde un activista de treinta años llamado Eli Pariser saldría al escenario para alertar al público, compuesto por ejecutivos tecnológicos e ingenieros, de que sus algoritmos podían poner en peligro la democracia. «Se está produciendo un cambio en cómo discurre la información por internet, y es invisible —dijo—. Y si no le prestamos atención, podría convertirse en un verdadero problema.»⁹

Un día, me contó Pariser, las publicaciones de sus amigos

conservadores habían desaparecido de su canal de noticias de Facebook y las publicaciones de sus amigos liberales habían empezado a aparecer de forma más destacada. Probablemente, el algoritmo de Facebook había detectado que Pariser interactuaba con mayor frecuencia con contenido liberal. No es ninguna sorpresa: es un activista progresista que durante años dirigió el portal de izquierdas MoveOn.org. Puede que el cambio hiciera aumentar el tiempo que pasaba en Facebook. Pero ¿era positivo para él que le mostrasen solo publicaciones que reforzaban los sesgos que ya tenía? ¿Era eso positivo para la sociedad? Pariser tenía un nombre para ese efecto: burbujas de opinión (*filter bubbles*, en inglés).

La clasificación algorítmica más simple puede alterar la actitud de los ciudadanos hasta el extremo de hacer decantar unas elecciones. En un experimento hecho en 2015, a unos ciudadanos estadounidenses se les pidió que escogieran entre dos candidatos ficticios investigándolos por internet.¹⁰ A cada participante le mostraron los mismos treinta resultados de búsqueda en una simulación de Google, pero en un orden distinto. De forma sistemática, los participantes dieron a los resultados situados más arriba un mayor peso psicológico, incluso aquellos que leyeron los treinta resultados. El efecto —concluyeron los investigadores— podía alterar hasta un 20 % la intención de voto de los participantes indecisos. El autor del estudio, Robert Epstein, psicólogo y fundador del Cambridge Center for Behavioral Studies, señaló en un artículo publicado en agosto de 2015 que «el próximo presidente de Estados Unidos podría llegar al cargo no solo gracias a los anuncios electorales televisados o a sus discursos, sino por las decisiones secretas de Google».¹¹ Observó que el extravagante y llamativo Donald Trump, a pesar de ser descartado por casi todos los medios de comunicación y las élites políticas incluso en el seno de su propio partido, estaba «ganando de paliza a todos los demás candidatos en cuanto a las búsquedas en cuarenta y siete de los cincuenta estados de Estados Unidos».

Varios años antes, el temor de Pariser había sido más básico. «Se está librando una batalla épica entre nuestro futuro, el yo que aspiramos a ser y nuestros yoes actuales, más impulsivos»,¹² declaró. Incluso en 2011, años antes de que YouTube o Facebook otorgaran superpoderes a sus sistemas para conseguir unos resultados tan

destructivos, esos algoritmos anteriores, más sencillos, ya tomaban partido de forma clara por los impulsos. Y solían ganar, lo que hacía proliferar «una autopropaganda invisible que nos adoctrina con nuestras propias ideas».

Al año siguiente, en 2012, Pariser fundó Upworthy, un portal web dedicado a contar noticias positivas, que publicaba listículos y vídeos con carga emocional diseñados para viajar por Facebook y YouTube. Pero, en lugar de aprovechar el poder del algoritmo para difundir contenido que promoviera el bien común, lo cual era la intención de Pariser, Upworthy quedó corrompido por esos sistemas, ya que sus creadores de vídeos usaban las preferencias del algoritmo para generar contenido que satisfacía las identidades y políticas de los usuarios sin enseñarles nada ni ilustrarlos, lo cual infestaba internet de comida basura informativa. Más adelante, Facebook modificó su algoritmo y el tráfico de Upworthy se evaporó.¹³ En realidad, no es que decenas de millones de personas hubieran elegido Upworthy, sino que se les había manipulado con máquinas para que leyeran su contenido.

A pesar de todo, a las compañías de redes sociales eso les preocupaba muy poco. En YouTube, a medida que iba actualizándose el algoritmo, algunas personas que había en las trincheras de la empresa, como Chaslot, terminaron por temer que el sistema estaba encerrando a los usuarios en peligrosas cámaras de resonancia de desinformación, de la misma forma en que el propio Chaslot había visto fomentar la misoginia viral entre los vídeos para los aficionados a los videojuegos. «No quería ser el típico francés que se queja todo el día», me dijo. «Quería aportar soluciones» con la creación de un algoritmo que atrajera a los usuarios satisfaciendo sus intereses y necesidades en lugar de explotar sus impulsos. «E intentaba hacerlo a la manera estadounidense —dijo—, centrándome en las oportunidades en vez de en los problemas.»

Sin embargo, en reiteradas ocasiones, sus jefes se negaron a escuchar las explicaciones sobre sus avances en un algoritmo alternativo, dijo Chaslot, o insistían en que nunca llegaría a ser una realidad. Él siguió trabajando en el proyecto de todos modos, dedicándole parte de su tiempo libre. Presentó los resultados a sus superiores, pero ellos volvieron a rechazarlo. Al final, empezó a trabajar en una tercera versión. «Mi director me dijo: “A ver, Guillaume, yo en tu lugar no haría este proyecto con tu 20 % de

tiempo libre porque tu rendimiento no es lo bastante bueno”.»

Chaslot lo apartó durante unos meses, pero el problema seguía reconcomiéndole. A lo largo del año 2013, los sistemas de YouTube parecían dar cada vez mayor prioridad a vídeos que eran, en el mejor de los casos, frívolos y adictivos («Bad Lip Readings») y, en el peor, llenos de odio y conspirativos, lo cual entrenaba a usuarios y creadores de contenido a seguirlos. Una de las voces más destacadas de la plataforma, el galán de quince años Nash Grier, había empezado el año colgando gags de broma y lo terminó, en lo que se estaba volviendo una trayectoria cada vez más habitual, con vídeos en los que aleccionaba a las mujeres sobre su deber tradicionalista de someterse a los hombres.¹⁴ YouTube estaba adiestrando a los usuarios a pasar el tiempo absorbiendo contenido que abarcaba desde la comida basura intelectual al veneno absoluto, algo muy alejado del viaje de iluminación y descubrimiento que Chaslot había creído que la plataforma hacía posible. «Es muy importante, tengo que seguir adelante con este proyecto —recordaba haber pensado—. Y luego me despidieron.»

YouTube mantiene que el despido de Chaslot, en octubre de ese año, se debió a su bajo rendimiento. Chaslot cree que fue despedido por denunciar algo que nadie quería oír. La diferencia, tal vez, sea mínima; YouTube se estaba reconfigurando en torno a un propósito único con el que Chaslot no estaba de acuerdo. «Esos valores de moderación, de amabilidad, cualquier cosa que puedas pensar que son valores en los que se basa nuestra sociedad, pues a los ingenieros no les importaba incorporar esos valores en el sistema —dijo—. Solo les importaban los ingresos publicitarios. Pensaban que, preocupándose de un único indicador, que es el tiempo de visualización, ya podías beneficiar a todo el mundo. Pero esto es sencillamente falso.»

MULTIPLICARSE POR DIEZ

Chaslot y otros herejes críticos con los algoritmos que le siguieron en los años posteriores estaban cuestionando algo más fundamental que los balances de cuentas de las empresas; estaban poniendo en cuestión su forma de ver el mundo. Desde la época de los semiconductores, los habitantes de Silicon Valley lo habían dado todo por los indicadores

cuantificables. Los productos o eran más eficientes o no lo eran. En los años ochenta, el director ejecutivo de Intel, Andy Grove, convirtió una variante extrema de esa idea en un dogma que se extendió por todo Silicon Valley, un dogma que había desarrollado mientras supervisaba la transición de la empresa de los circuitos integrados —los intestinos electrónicos, la culminación de la evolución de los semiconductores— a los microprocesadores, los cerebros de los productos digitales.

Silicon Valley estaba perdiendo su dominio en el mercado de los semiconductores para los circuitos integrados ante Japón y Taiwán. Los microprocesadores podían ser una opción de futuro para los fabricantes de chips como Intel, pero había un problema. El predecesor de Grove en Intel, un ingeniero de toda la vida llamado Gordon Moore, había acuñado lo que acabó conociéndose como la ley de Moore, según la cual la capacidad de procesamiento se doblaría cada dos años. Un crecimiento exponencial tan constante, imposible en cualquier otro sector, daría lugar a unas posibilidades transformadoras. Imaginad que los aviones de pasajeros doblaran su velocidad cada veinticuatro meses o que las baterías pudieran almacenar el doble de carga en ese periodo. Eso significaba que la cifra se multiplicaría por 32 en diez años y por 1.024 en veinte años. Pero mantener el ritmo exigía hacer avanzar la tecnología y su producción a ese ritmo vertiginoso. No estaba claro que tal cosa fuera posible. ¿Las empresas asiáticas también dominarían ese mercado? Grove reorientó todos los aspectos del negocio alrededor de unos pocos indicadores, como la velocidad de los procesadores o el tiempo de comercialización, y concedió atribuciones a ingenieros de a pie para que hicieran todo lo que hiciese falta para alcanzar los objetivos. Así, los legendarios fabricantes de chips de Silicon Valley no solo sobrevivieron, sino que aquel mercado, dominado por Intel e impulsado por la ley de Moore, se disparó mientras el resto de la economía estadounidense se hundía.

Cuando Grove se vio forzado a renunciar al cargo al diagnosticársele un cáncer en el momento cumbre de su éxito, se pasó el periodo de convalecencia escribiendo libros y dando conferencias. Como un evangelista errante en los desesperados días de la crisis de las puntocom, encontró muchos discípulos entusiastas. Dedicad todos los esfuerzos, los instruía, a maximizar unos pocos indicadores cuantificables. Concentrad el poder en manos de unos ingenieros que

puedan hacerlo. Y relegad al resto. Entre sus seguidores estaba John Doerr, un comercial de Intel convertido en un influyente inversor de capital riesgo, que inculcó la filosofía de Grove, obsesionada en unos indicadores, a montones de empresas primerizas de internet. Una de ellas era Amazon, cuyo fundador, Jeff Bezos, escribió lo siguiente en una carta a los accionistas que aún circula por Silicon Valley: «Hay una respuesta correcta o una respuesta incorrecta; una respuesta mejor o una peor; y las matemáticas nos dicen cuál es cuál». ¹⁵ Otra era Google, a cuyos jóvenes fundadores Doerr instruyó personalmente en el evangelio de Grove. Wojcicki asistió como oyente.

No obstante, a medida que Silicon Valley fue ampliando su alcance, esa cultura de la optimización a cualquier precio conllevó efectos secundarios. Al optimizar para obtener los trayectos compartidos más rápidos, Uber eliminó las protecciones laborales del mercado internacional del taxi. Al optimizar los ingresos por los alquileres a corto plazo, Airbnb hizo que las viviendas para el largo plazo fueran más escasas y se encarecieran. Las redes sociales, al optimizar para obtener la mayor cantidad posible de usuarios y el mayor tiempo posible de actividad en las plataformas, tal vez fueron lo que tuvo un mayor impacto de todo. «Era una forma fantástica de crear una *start-up* —dijo Chaslot—. Te fijas en un solo indicador y todo el mundo trabaja para maximizar ese indicador. Y eso es muy eficaz para obtener crecimiento. Pero es un desastre para otras muchas cosas.»

En la mayor parte de los negocios, los indicadores podían crecer entre un 3 y un 4 % anuales. Pero Grove, interiorizando la ley de Moore, había insistido en que las empresas encontrasen «cambios que multipliquen por diez»: innovaciones que hagan avanzar la empresa por un factor de diez. Justo el mismo factor que Goodrow, en 2012, había aceptado como objetivo con respecto al tiempo de visualización de YouTube. «Los mil millones de horas diarias se habían convertido en su ballena blanca», escribió Wojcicki, la ejecutiva de Google, y no solo para Goodrow. «El OKR de los mil millones de horas era una religión en YouTube —observó Wojcicki empleando un acrónimo empresarial para referirse a los indicadores (*Objectives and Key Results*, ‘objetivos y resultados clave’)—, hasta el punto de excluir casi todos los demás elementos.» ¹⁶

Ese mismo año, Renée DiResta, la inversora tecnológica que más

adelante rastrearía antivacunas en Facebook en sus ratos libres, se dio cuenta de que el mandato de «multiplicar por diez» de Grove se estaba transformando en un nuevo y extraño modelo de negocio muy distinto del que había dado lugar a empresas como Intel. La primera vez que había visto ese cambio había sido en uno de los importantísimos congresos de inversión de Silicon Valley, presentado por una aceleradora de *start-ups* tecnológicas llamada Y Combinator, donde los creadores de empresas se ponían en contacto con intermediarios financieros que podrían conseguirles capital. «El *Demo Day* de YC era como recibir una invitación para asistir a los Óscar», afirmó. La jornada anual de presentación de las empresas salidas de la incubadora de Y Combinator «no era un evento en el que pudiera entrar cualquiera». Los inversores de bajo rango como DiResta, sin capacidad de extender un cheque *in situ*, no eran bienvenidos, pero su jefe, como había perdido el avión, le pidió que fuera ella en su lugar.

Cuando empezaron las presentaciones, a cada empresario incipiente se le concedió dos minutos para convencer al público, lleno de inversores de peso. Las ideas —un servicio de computación en la nube, un portal de inversiones, una web para reservar viajes o un agregador de memes— eran muy variadas, pero todas tenían el mismo plan de negocio. «Enseñaban un gráfico en el que se veía como su trayectoria iba hacia arriba y hacia a la derecha —recordaba DiResta—. Casi ninguna sufría bajadas. No era más que un garabato que iba hacia arriba y hacia la derecha. Yo lo veía y pensaba: “Pero ¿qué coño está pasando?”.»

DiResta estaba especializada en tecnología del *hardware*, igual que Intel. Las empresas tenían que invertir en gastos de estructura, planificar la fabricación y la logística e identificar clientes, todo ello antes de distribuir la primera unidad. Requería hacer proyecciones de costes y ventas con enorme detalle. Las presentaciones de aquel congreso eran mucho más vagas: las empresas diseñarían una página web, mucha gente la usaría, venderían anuncios y todo el mundo se haría rico. A nadie del público parecía importarle la falta de detalles concretos; DiResta oyó rumores de valoraciones de 15 o 20 millones de dólares.

Durante los siguientes meses, vio como una *start-up* tras otra obtenían inversiones de seis o siete cifras a pesar de tener el mismo plan de negocio «simplista, hecho para una presentación de clase» y a

menudo sin ingresos. Los inversores, se percató DiResta, no inyectaban dinero a cualquier chaval con buen discurso. Buscaban un modelo muy concreto: servicios web de uso gratuito que prometían un crecimiento meteórico del número de usuarios. No obstante, eso la desconcertó, porque muchas cerraban sin obtener beneficios y acto seguido llegaba otra ronda de *start-ups* que las sustituía. «No sabía si era demasiado escéptica o no era lo bastante visionaria», admitió. ¿Qué se le estaba escapando?

La respuesta resultó ser una tecnología preponderante en los últimos tiempos llamada computación en la nube. Antes de la computación en la nube, si querías poner en marcha un negocio virtual, tenías que invertir en servidores y en todas las infraestructuras que los rodeaban: espacio de oficina, banda ancha, control especial de la climatización y personal para supervisarlos todo. Eso podía suponer una inversión inicial millonaria. Requería convencer a un inversor de capital riesgo de que recuperaría su inversión, por lo general prometiendo vender bienes o servicios, lo cual añadía aún más gastos de estructura. Era una situación que hacía que los inversores fuesen conservadores. Una apuesta fracasada de 20 millones de dólares podía ser devastadora e, incluso si la empresa tenía éxito, se podía tardar años en generar beneficios.

Luego, a finales de la primera década del siglo XXI, Amazon y algunas empresas más crearon enormes granjas de servidores y empezaron a ofrecer su capacidad de procesamiento y de almacenamiento de datos en alquiler. A eso lo llamaron «la nube». Ahora ya no tenías que invertir en gastos de estructura. Podías conseguir tus primeros 10.000 clientes pidiendo un préstamo a papá y mamá. No necesitabas un modelo de rentabilidad, inversores iniciales ni siquiera una idea concebida al detalle. «Olvidaos de las estrategias —escribió el inversor Roger McNamee refiriéndose a ese nuevo enfoque—. Juntad a unos cuantos amigos, cread un producto que os guste e intentad comercializarlo. Cometed errores, corregidlos y volved a empezar el ciclo.»¹⁷ Era transformador también para los inversores, quienes ya no tenían que dejarse millones para introducir una *start-up* en el mercado. Podían hacerlo con calderilla.

Eso cambió lo que los inversores querían sacar de sus inversiones. Ya no se trataba de encontrar ese fabricante de artilugios prometedor cuyas ventas algún día, tras muchos años de empeño e inversiones,

podrían eclipsar los costes. Se trataba de invertir en muchas *start-ups* de internet baratas sabiendo que la mayoría fracasarían, pero que si una tenía éxito cubriría esas pérdidas e incluso más.

Sin embargo, también se estaba poniendo patas arriba la definición de éxito. Las grandes compañías estaban cada vez más desesperadas por encontrar un lugar en el sector de internet adquiriendo empresas emergentes por cantidades apabullantes. Y los corredores de bolsa de Wall Street, deseosos de comprar acciones del próximo gran éxito tecnológico, regaban con capital cualquier cosa que se pareciese, aunque fuera vagamente, al Microsoft o al Apple del mañana. En 1994, Kleiner Perkins aportó cinco millones para adquirir el 25 % de una *start-up* llamada Netscape, que había creado un navegador web. La empresa no tenía beneficios, en parte porque ofrecía su producto gratis, aunque eso le hizo ganar millones de usuarios. Al año siguiente, Netscape empezó a cotizar en bolsa. Los especuladores bursátiles hicieron subir en poco tiempo su valoración hasta los 2.300 millones de dólares, lo cual reportó a Kleiner un rendimiento que multiplicaba por cien su inversión. Netscape desapareció al cabo de unos años.

A medida que la computación en la nube permitió la proliferación de empresas emergentes del estilo de Netscape, el incentivo predominante para los inversores tecnológicos pasó a ser preparar a empresas emergentes para que se consumieran con rapidez e intensidad, lo cual les garantizaba una venta cuantiosa al poco tiempo o una oferta pública inicial. Y la mejor manera de sacar mucho dinero gastando poco es invertir en servicios web que no ofrezcan productos físicos pero que atraigan a muchísimos usuarios. Así, parece que algún día tu inversión podría convertir a esa base de usuarios en una fuente de beneficios, lo cual da lugar, aunque sea algo totalmente teórico, a una elevada valoración de mercado.

El gráfico del garabato ascendente —tal como lo veía DiResta—, con su promesa de utilizar unos servicios digitales gratuitos para atraer el mayor número de usuarios posible sin un plan de rentabilidad, no era ninguna trampa. Era lo que exigían los inversores. Y, en ese modelo, nada funcionaba mejor que las redes sociales. «Tal vez uno vería esas cifras y llegaría a la conclusión de que los inversores se habían vuelto locos», escribió una vez Peter Thiel. Pero la recompensa podía ser astronómica. Si alguien invirtió 250.000

dólares en Instagram en 2010, ese dinero aumentó a 78 millones de dólares dos años después, cuando Facebook adquirió la empresa. Si los planes de negocio de los creadores eran tan absurdos como le habían parecido a DiResta, no pasaba nada. Las pérdidas eran bajas, mientras que un éxito, aunque fuera infrecuente, haría rico a todo el mundo. Era el culto del gráfico del garabato.

«Me daba la sensación de que, en algunos momentos, en la financiación de las *start-ups* podía haber un rollo tipo “el emperador va desnudo” —me dijo DiResta—. ¿Hasta cuándo se podía perpetuar esa valoración descabellada antes de que se llegara al punto en que alguien tuviera que extender un cheque y se produjera la oferta pública inicial, y se pidiera a personas normales y corrientes que invirtiesen sus ahorros para la jubilación en una empresa social-móvil-local con cupones? Por eso aún no soy inversora en capital riesgo.»

Aunque la empresa emergente sobreviviera, tenía el lastre de una valoración que se había inflado hasta unos extremos desmesurados por culpa de todas aquellas inversiones especulativas. Para impedir que el precio de sus acciones se desplomase (y mantener el negocio en funcionamiento, una vez que el dinero de los inversos se hubiera agotado), su única opción era convertir a todos aquellos usuarios en dinero vendiéndoles anuncios. El valor de la publicidad es la atención: tu ojo echa un vistazo a un *banner* publicitario, que Facebook vende a Toyota o a Gap más o menos por un céntimo. Pero el presupuesto publicitario de Toyota es fijo. Así como la atención humana total. Por tanto, cada vez que una red social actualiza sus sistemas para robarle unos cuantos minutos más a una persona, da un paso más en una carrera armamentística tecnológica para acaparar tu campo visual. Y a medida que se incrementa el suministro de anuncios virtuales, baja el precio. En un memorando de 2014, el director ejecutivo de Microsoft aseguró que «la verdadera materia prima que escasea es cada vez más la atención de los seres humanos».¹⁸

Si el valor de la impresión dejada por un anuncio seguía menguando, incluso plataformas como Facebook y YouTube podían dejar de ser viables. Su única elección era aumentar permanentemente el número de usuarios y el tiempo que dichos usuarios pasaban en la plataforma, muchas veces más deprisa de lo que esas mismas acciones hacían bajar el precio de un anuncio. Pero controlar el mercado de la atención humana, como sus modelos de negocio los habían destinado

a intentar, estaba fuera del alcance de cualquier cosa que un programa hecho por el hombre pudiera lograr.

Las empresas, para sobrevivir a ese entorno creado por ellas mismas, tendrían que confiar sus negocios —y, por tanto, a sus usuarios— a las máquinas.

PATRONES TURBIOS

En 2014, Wojcicki, la ejecutiva de Google que había recomendado que se adquiriese YouTube, asumió el cargo de directora ejecutiva del servicio de vídeo. Aunque su liderazgo estaría obsesionado con el crecimiento de una forma tan implacable como el de Mark Zuckerberg o el de Jack Dorsey, ella solo atraería una pequeña parte de las críticas. Incluso en el apogeo de la respuesta negativa ante las redes sociales —cuando se acusó de forma convincente a su plataforma de provocar perjuicios aún mayores que los de Facebook—, ella casi nunca tuvo que comparecer ante el Congreso estadounidense, casi nunca fue regañada por presentadores de telediario de cadenas por cable y casi nunca se mencionó su nombre. Su historial, más convencional —había trabajado en departamentos de publicidad y *marketing* y no había sido una prodigiosa *hacker* de residencia estudiantil—, le quitó las ganas de convertirse en el personaje superestrella del mundo tecnológico —y, sobre todo, de las declaraciones revolucionarias— que granjearon a los Zuckerbergs y Dorseys tanta adoración que terminaron metiéndoles en un lío. Por otra parte, Google trató YouTube como su máquina de liquidez en vez de como su marca líder, lo que dejó el servicio y a Wojcicki en un segundo plano.

Más importante aún, quizás, fue que YouTube nunca compartió las pretensiones de Facebook, Twitter o Reddit de salvar el mundo, y en años posteriores casi nunca siguió las iniciativas de esas compañías de demostrar en público que estaban repensando su lugar en el funcionamiento de la sociedad. El YouTube de Wojcicki existía para convertir visualizaciones en dinero. La democracia y la cohesión social no eran problema suyo.

Poco después de que Wojcicki accediese al cargo, Goodrow la alertó: «No vamos a lograr ese OKR de tiempo de visualización si no

hacemos algo».¹⁹ ¿Y qué era ese algo?: dar un mayor poder a unas inteligencias artificiales cada vez más inescrutables. En un artículo publicado en 2016, los ingenieros de Google anunciaron un «cambio de paradigma radical» hacia una nueva clase de aprendizaje automático que denominaron «aprendizaje profundo».²⁰

En la IA anterior, un sistema automatizado creaba los programas que seleccionaban los vídeos. Pero, como en el caso de las IA que detectaban correo basura, los humanos supervisaban ese sistema y, durante su evolución, intervenían para guiarlo y hacer cambios. Ahora, el aprendizaje profundo era lo bastante sofisticado para encargarse también de la supervisión. Por consiguiente, en la mayoría de los casos «no habrá humanos haciendo los ajustes de los algoritmos, ni midiendo esos ajustes ni luego implementándolos», escribió el jefe de una agencia que creaba talento para YouTube en un artículo en el que desgranaba el documento sobre el aprendizaje profundo. «Así pues, cuando YouTube afirma que no saben con exactitud por qué el algoritmo hace lo que hace, probablemente lo dicen de una forma muy literal.»²¹

Era como si Coca-Cola llenase mil millones de máquinas de refrescos con una bebida diseñada por una IA sin que un solo ser humano evaluara el contenido de las botellas. Y si la IA encargada del proceso estuviera programada solo para hacer subir las ventas, sin tener en cuenta la salud ni la seguridad. Como dijo uno de los ingenieros encargados del aprendizaje automático de YouTube en un congreso sectorial, «el departamento de producción nos dice que queremos incrementar tal indicador, y luego vamos nosotros y lo incrementamos».²²

El tiempo medio que dedicaban los usuarios a la plataforma se disparó. La compañía calculó que un 70 % del tiempo pasado en la plataforma —un porcentaje astronómico de su negocio— era resultado de que los vídeos aparecieran según el sistema de recomendaciones gestionado por el algoritmo.²³

A medida que avanzó esa tecnología, otras plataformas también ampliaron el uso de los algoritmos autodirigidos: Facebook para seleccionar qué publicaciones ven los usuarios y a qué grupos se les recomienda unirse; Twitter para dar mayor prioridad a las publicaciones que pudieran tentar a un usuario a seguir desplazándose hacia abajo y a tuitear.

«Diseñamos muchos algoritmos para producir contenido interesante para vosotros —dijo Zuckerberg en una entrevista—. El algoritmo analiza toda la información disponible de cada usuario y, de hecho, calcula cuál va a ser la información más interesante.»²⁴ Un extrabajador de Facebook lo dijo con menos rodeos: «Está diseñado para hacer que quieras seguir desplazándote página abajo, seguir mirando y poniendo “me gusta”».²⁵ Y otro: «Esa es la clave. Ese es el ingrediente secreto. Ahí están el cómo y el porqué valemos tantos miles de millones de dólares».²⁶

En 2014, el mismo año en que Wojcicki asumió la dirección de YouTube, el algoritmo de Facebook sustituyó su preferencia por ciberanzuelos del estilo de Upworthy por algo todavía más magnético: interacciones con implicación emocional. A lo largo de la segunda mitad de ese año, mientras la empresa iba reconfigurando sus sistemas, los investigadores internos de la plataforma rastrearon diez millones de usuarios para entender los efectos.²⁷ Hallaron que los cambios inflaban de forma artificial la cantidad de contenido proliberal que veían los usuarios liberales y la cantidad de contenido proconservador que veían los conservadores. Aquello de lo que Pariser había alertado. El resultado, aunque nadie de Facebook hubiese tenido la intención consciente de hacerlo, fue una exacerbación del partidismo introducida por el algoritmo. Eso era más poderoso que clasificar a las personas en el equivalente en Facebook de un canal de noticias de Fox News o la MSNBC, porque, mientras que la relación entre una cadena de televisión por cable y el espectador es unidireccional, la relación entre el algoritmo de Facebook y el usuario es bidireccional. Uno entrena al otro y viceversa. El proceso —como explicaron con cierta cautela los investigadores de Facebook en un tácito aviso al que la empresa no hizo caso— estaba «asociado con la adopción, a lo largo del tiempo, de actitudes más extremas y con la percepción distorsionada de las noticias de actualidad».²⁸

Pero las ambiciones algorítmicas de Silicon Valley no hicieron más que crecer, hasta nada menos que el dominio de la mente humana. Durante un acto corporativo celebrado el verano siguiente con Zuckerberg y Stephen Hawking, el físico le preguntó al jefe de Facebook:²⁹

—¿De cuál de las preguntas de la ciencia te gustaría saber la respuesta y por qué?

Zuckerberg respondió:

—Las que más me interesan son las cuestiones relacionadas con las personas. También siento curiosidad por saber si hay una ley matemática fundamental que subyazca a las relaciones sociales entre humanos y que rijan el equilibrio entre qué personas y qué cosas nos importan a todos. Yo apuesto a que hay una.

Facebook contrató a dos de los principales expertos mundiales en aprendizaje automático, estrellas del rock por méritos propios, para dirigir laboratorios internos de IA. La compañía tenía la intención no solo de aprovechar los avances hechos en este ámbito, sino también de impulsar que hubiera más. «Cada vez que utilizas Facebook o Instagram o Messenger —dijo uno de los directores de los laboratorios en un congreso sectorial—, puede que no te des cuenta, pero tus experiencias están siendo impulsadas por la IA.»³⁰ Si Facebook quería que miles de millones de usuarios hicieran algo en mayor o menor medida —clicar, poner «me gusta» o comentar—, lo único que tenía que hacer era pedir a su sistema que lo hiciera realidad.

Si las compañías no se molestaban en comprobar cómo el sistema lograba eso, era porque albergaban el mismo optimismo cegador que Eric Schmidt, director ejecutivo de Google durante años, había manifestado en Birmania: más presencia en la plataforma significaba un mayor bien común. No nos preocupemos de comprobar si eso es verdad, que tal vez descubriríamos que no lo es. «Si hacen esas interacciones un poco más, eso indica que estamos creando valor —dijo en una ocasión en un congreso de Facebook Adam Mosseri, el vicepresidente que entonces supervisaba el canal de noticias—. A largo plazo, más y más personas utilizarán Facebook, y van a pasar más y más tiempo en la plataforma, y eso va a ser bueno para ellos, para Facebook y para las empresas editoriales.»³¹ Esa concepción era generalizada. Goodrow, el responsable de algoritmos de YouTube, había escrito que «cuando los usuarios dedican una mayor parte de su valioso tiempo a ver vídeos de YouTube, deben por fuerza ser más felices con esos vídeos».³²

Era una suposición extraña. Las personas actúan por sistema contra sus propios intereses. Bebemos y comemos en exceso, consumimos drogas peligrosas, procrastinamos, sucumbimos a las tentaciones del narcisismo o el odio. Perdemos los estribos, el autocontrol y nos olvidamos de nuestros principios morales. Existen

disciplinas enteras basadas en la idea de que nuestros impulsos pueden imponerse a nosotros, a menudo en detrimento de nosotros mismos. Esa era una verdad capital de la experiencia humana, incompatible con unas economías de crecimiento exponencial; y, por tanto, esa idea se desoyó oportunamente.

El deseo de aumentar la participación de los usuarios, que ha seguido siendo una característica permanente de la cultura empresarial, es totalizador. Los ingenieros de Facebook eran «llamados a capítulo» de forma automática —recordaba un exdirector del equipo encargado del canal de noticias— si disminuían los «me gusta» o las comparticiones, para que pudieran reajustar el sistema y hacer que volvieran a subir. «Si tu trabajo consiste en hacer que esa cifra suba, llega un punto en el que se te acaban los métodos buenos y totalmente positivos³³—ha dicho un exjefe de operaciones de Facebook—. Y empiezas a pensar: “A ver, ¿cuáles son los patrones turbios que puedo emplear para conseguir que la gente vuelva a conectarse a la plataforma?”.»³⁴

Las empresas aprendieron a quitarle importancia al grado en que los robots diseñaban la realidad de miles de millones de personas. Pero en algunas ocasiones se han filtrado pistas del poder de las máquinas. La aplicación china TikTok muestra a cada usuario un hilo de vídeos seleccionado casi por completo por algoritmos. Su IA es tan sofisticada que TikTok atrajo casi al instante 80 millones de usuarios estadounidenses, quienes a menudo utilizan la plataforma durante horas, a pesar de que la mayor parte de sus ingenieros ni hablan inglés ni entienden la cultura estadounidense.

«Un algoritmo de aprendizaje automático con una sensibilidad y precisión significativas puede penetrar el velo de la ignorancia cultural —escribió el inversor Eugene Wei en referencia a TikTok—. La cultura puede abstraerse.»³⁵O, como afirmó un ingeniero del equipo responsable de los algoritmos de YouTube en declaraciones al *Wall Street Journal*, «no tenemos que pensar tanto». Puesto que el algoritmo es quien se encarga de las cosas, «le damos datos sin procesar y dejamos que se apañe».³⁶

Al empezar el año 2015, DiResta ya llevaba casi un año rastreando las consecuencias de esos algoritmos. Sus esfuerzos para entender las redes de antivacunas —y los de los demás analistas de las redes sociales que había conocido a raíz de ello— culminaron, ese

verano, en unas reuniones en la sede central del Departamento de Estado en Washington. El Gobierno, que combatía a Estado Islámico sobre el terreno y en internet, se había percatado de que los sistemas promocionales de las redes sociales fomentaban a los yihadistas o incluso exageraban su alcance. DiResta y sus compañeros facebookólogos aportaron sus conocimientos sobre cómo funcionaban las plataformas. A lo largo de los siguientes meses, se expulsaría en su mayor parte a los yihadistas de la web social. Pero las reuniones pusieron de manifiesto dos importantes lecciones para DiResta. Primero, que ella y sus compañeros friquis de la informática ya no eran los únicos que estaban preocupados por los peligros de las redes sociales. En Washington había altos funcionarios que compartían sus preocupaciones. Y segundo, que los analistas del Gobierno presentes en la reunión no dejaron de llamar la atención sobre una amenaza virtual que DiResta y sus compañeros aún no habían encontrado: los servicios de inteligencia rusos.

Al igual que los antivacunas de DiResta o incluso como Upworthy, los rusos explotaban las propias preferencias del algoritmo. No era solo que los agentes repitiesen expresiones o comportamientos que daban un buen rendimiento. Su aparente misión —generar conflicto político— parecía alinearse de forma natural con lo que los algoritmos ya favorecían, a menudo hasta los extremos. La polémica, el tribalismo y las conspiraciones. Pero la facilidad de los rusos a la hora de explotar eso era, según concluyó DiResta, un síntoma; el problema era el sistema. Invitaba a esa manipulación. Incluso la recompensaba.

AVISOS DE ALERTA

Tras ser despedido, Guillaume Chaslot regresó a Francia. Se pasó un par de años en un portal de comercio electrónico francés. Silicon Valley era un recuerdo lejano. Hasta que, en un largo trayecto en autobús que hizo a finales de 2015, le llamó la atención el móvil del pasajero que tenía al lado. El hombre estaba viendo YouTube, y todos los vídeos eran sobre conspiraciones. Lo primero que pensó Chaslot fue lo que pensaría un ingeniero: «Qué sesión audiovisual tan fantástica...». El algoritmo de recomendación de vídeos saltaba de un

tema a otro, siempre con experiencias nuevas, a la vez que hundía al hombre más y más en el abismo. «Entonces me di cuenta —dijo Chaslot— de que, desde un punto de vista humano, eso es, de hecho, una catástrofe. El algoritmo que yo había contribuido a elaborar estaba llevando a ese hombre hacia vídeos más y más llenos de odio.»

Chaslot entabló conversación con aquel pasajero y le preguntó por el vídeo que estaba viendo en ese instante, en el que se describía una trama para exterminar a miles de millones de personas. Tenía la esperanza de que el hombre soltase una carcajada como muestra de que sabía que era algo absurdo. Pero le dijo a Chaslot: «Tienes que verlo». Le explicó que los medios de comunicación nunca contaban esos secretos, pero que la verdad estaba ahí, en YouTube. «No puedes creer todo lo que hay por internet», le dijo Chaslot. Pero estaba demasiado avergonzado para admitir ante aquel hombre que había trabajado en YouTube y que por eso sabía que su sistema lanzaba a los usuarios por madrigueras sin ningún respeto por la verdad. «El hombre me decía: “Es que hay muchos vídeos, tiene que ser cierto” — me dijo Chaslot—. Lo que le convencía no eran los vídeos por separado, sino la repetición. Y la repetición se debía al motor de recomendaciones.»

YouTube estaba explotando un resquicio cognitivo conocido como el efecto de la verdad ilusoria. Todas las horas del día recibimos un bombardeo de información. Para sobrellevar eso, tomamos atajos mentales para decidir con rapidez qué aceptar y qué rechazar. Uno es la familiaridad; si una afirmación nos suena como algo que antes ya hemos aceptado como verdad, probablemente siga siéndolo. Es un hueco que tenemos en nuestras defensas mentales por el que podría pasar un camión. En varios experimentos, a los participantes a quienes se les ha bombardeado con la expresión «la temperatura corporal de un pollo» aceptarán como válidas sin reparos variaciones como «la temperatura corporal de los pollos es de 62 °C».³⁷ El vecino de asiento de Chaslot había estado expuesto a las mismas conspiraciones descabelladas tantas veces que su mente probablemente confundía la familiaridad con el olorcillo a verdad. Como ocurre con el resto de cosas en las redes sociales, el efecto se agudiza debido a una falsa sensación de consenso social, lo cual desata nuestros instintos de aquiescencia.

Chaslot había oído historias de personas que se habían caído por

madrigueras youtuberas. Pero la convicción que emanaba de la voz de ese hombre, que por lo demás tenía un aspecto normal, le preocupó. ¿Acaso otras personas eran víctimas de lo mismo? Creó un programa sencillo, al que llamó Algo Transparency, para averiguarlo. El programa escribía un término —por ejemplo, el nombre de un político— en el recuadro de búsqueda de YouTube. Luego abría los resultados que salían más arriba. Acto seguido, cada recomendación de qué ver a continuación. Hizo enormes tandas de búsquedas anónimas, una tras otra, durante finales de 2015 y buena parte de 2016, para detectar tendencias.

Lo que encontró le alarmó. Cuando buscó «papa Francisco» en YouTube, por ejemplo, un 10 % de los vídeos que se mostraban eran conspiraciones.³⁸ Sobre «cambio climático», era de un 15 %. Pero el verdadero *shock* lo tuvo cuando siguió las recomendaciones del algoritmo sobre qué ver a continuación, que YouTube ha afirmado que representa la mayor parte del tiempo de visualización en la plataforma. Un asombroso 85 % de los vídeos recomendados sobre el papa Francisco eran conspiraciones, que por ejemplo afirmaban la «verdadera» identidad del pontífice o decían poner al descubierto complots satánicos en el Vaticano. Sobre el calentamiento global, la cifra era de un 70 %, y por lo general lo calificaban de engaño. Sobre cuestiones con menos conspiraciones consolidadas, el sistema parecía hacerlas aparecer de la nada. Cuando Chaslot buscó «Quién es Michelle Obama», por ejemplo, poco menos de la mitad de los resultados principales y casi dos tercios de las recomendaciones sobre qué ver a continuación afirmaban que la primera dama era en realidad un hombre. Entonces pensó que, aunque él y sus antiguos compañeros de trabajo no estuvieran de acuerdo, seguro que querrían conocer aquello. Pero cuando contó sus preocupaciones en privado a personas que conocía en YouTube, la respuesta fue siempre la misma: «Si la gente clica ese contenido perjudicial, ¿quiénes somos nosotros para juzgarlos?».

Dentro de Google, no obstante, algunas personas estaban llegando a conclusiones parecidas a las de Chaslot. En 2013, un ingeniero llamado Tristan Harris difundió una circular urgiendo a la compañía a plantearse el impacto en la sociedad de las notificaciones con recuadros y vibraciones que buscaban la atención de los usuarios.³⁹ Como había sido alumno del Persuasive Tech Lab de la

Universidad de Stanford, conocía el poder de la manipulación. ¿Todo ese entrenamiento cognitivo podía tener un precio? Le otorgaron el título de «moralista del diseño», pero le dieron poco poder y, en 2015, dejó la empresa con la esperanza de presionar al sector para que cambiara. En una presentación que dio ese año en Facebook, Harris citó pruebas de que las redes sociales provocaban sentimientos de soledad y alienación, y se refirió a ello como una oportunidad para frenar el efecto. «No hicieron nada para resolverlo —explicó a *The New Yorker*—. Mis argumentos les quedaban en un punto ciego.»⁴⁰ Se desplazó por Silicon Valley alertando de que sus IA, un ejército de robots dedicado a vencer el control que cada usuario tiene de su propia atención, estaban librando una guerra invisible contra miles de millones de consumidores.

Otro empleado de Google, James Williams —quien tiempo después escribió ensayos en los que afirmaba que el Gamergate había sido una señal de alerta de que las redes sociales darían alas a Trump —, tuvo su epifanía mientras monitorizaba un *dashboard* que rastreaba las interacciones en tiempo real de los usuarios con anuncios. «Me di cuenta de una cosa: de alguna forma hemos incitado o convencido a un millón de personas de que hagan algo que, de lo contrario, no harían»,⁴¹ ha dicho. Se sumó a los esfuerzos de Harris dentro de Google hasta que, al igual que Harris, dejó el trabajo. Pero, en lugar de convencer a Silicon Valley, intentó hacer sonar las alarmas en la sociedad. «No existe una buena analogía para describir este monopolio de la mente que ahora tienen las fuerzas de la persuasión industrializada», escribió. El mundo se encontraba ante «una amenaza de última generación a la libertad humana» que se había «materializado justo delante de nuestras narices».⁴²

A lo largo de 2016 también sonaron algunas señales de alerta parecidas a estas, y no venían de luditas mal informados ni de activistas movidos por un programa político. Las advertencias vinieron de personas internas de Silicon Valley que conocían la tecnología y las plataformas, que compartían los ideales y los supuestos de ese lugar. Una de ellas era DiResta. «Estoy muy segura de que vuestro motor de recomendaciones está arrastrando a la gente a ese contenido», dijo a contactos que tenía en las compañías con la esperanza de que utilizarían sus herramientas internas de recopilación de datos para investigar el aumento repentino de la desinformación política y los

rumores polarizadores.

Ese verano, algunas personas dentro de Google que habían oído hablar del trabajo de DiResta sobre el rastreo de grupos antivacunas le pidieron que diera una conferencia en el congreso anual de la compañía. Hasta entonces, DiResta solo había mencionado sus preocupaciones más generales por canales privados, pero entonces decidió hacerlo público y advertir a todo un auditorio de ingenieros y directivos de alto rango de que sus productos representaban un peligro creciente para la sociedad. Paseándose por el escenario con un micrófono inalámbrico, embarazada de cuatro meses de su segundo hijo, DiResta enumeró conspiraciones típicas en internet: que los ciudadanos de Oregón temían que el agua contuviera fluoruro o que los brasileños pensaban que el zika era una trama malvada pero que «no estaban seguros de si culpar a las vacunas, a los transgénicos, a las estelas químicas o a Monsanto».⁴³

El público fue riendo con sus explicaciones, recreándose mientras se burlaban de aquellos ingenuos catetos. Hasta que llegó al remate final del chiste: en el fondo, los algoritmos de las redes sociales — también los de YouTube, propiedad de Google— eran responsables de aquello. «Hemos llegado a un punto en que es mucho más probable que veas cosas que son populares y que te tocan emocionalmente a que veas cosas que son verdad», les espetó.

Además, lejos de ser un fenómeno marginal, esas conspiraciones representaban un cambio más drástico que había sido provocado por las redes sociales. Los habitantes de Oregón habían votado a favor de prohibir el fluoruro y los brasileños habían dado marcha atrás en las protecciones contra el zika: en ambos casos, apoyándose en creencias que habían madurado en internet. «Los algoritmos están influyendo en las decisiones políticas», afirmó DiResta. El problema iba a empeorar a menos que se reformulasen con «un sentido de responsabilidad cívica». Sus palabras fueron recibidas con educación pero con escasa efusividad. «Lo que me dijeron sobre todo fue: “Qué teoría tan interesante”», recordaba DiResta.⁴⁴

Sin embargo, la influencia de los algoritmos no hizo sino agudizarse, también en el último reducto, Twitter. Durante años, la plataforma había mostrado al usuario un canal de noticias sencillo y cronológico de los tuits de sus amigos. Hasta que, en 2016, introdujo un algoritmo que clasificaba las publicaciones: para incrementar la

participación de los usuarios, por supuesto, y con unos efectos predecibles. «El tuit medio seleccionado era más emotivo, en todos los sentidos, que su equivalente cronológico», descubrió *The Economist* en un análisis de la modificación. El resultado fue justo lo que había sucedido en Facebook y YouTube: «El motor de recomendación parece priorizar el lenguaje incendiario y las afirmaciones disparatadas». ⁴⁵

Para los usuarios, para quienes el algoritmo era invisible, aquellos tuits funcionaban como unos potentes códigos sociales. Era como si de repente tu comunidad hubiese decidido que valoraba la provocación y la indignación por encima de todo, y recompensaba esas actitudes con oleadas de atención que, en realidad, estaban generadas por el algoritmo. Y, como el algoritmo daba menor prioridad a las publicaciones que consideraba contrarias a la participación, también podía decirse lo contrario. Daba la sensación de que de repente tus contactos despreciasen los matices y la moderación emocional y que de forma implícita te rechazaran ignorándote. Los usuarios parecieron absorber esas señales y se volvieron más crueles y se enfadaban más, dispuestos a humillar a miembros del exogrupo, a castigar a los transgresores sociales y a validar la visión del mundo de quienes pensaban igual.

El director ejecutivo de Twitter, Jack Dorsey, reconoció más adelante que, en una plataforma optimizada con un algoritmo para maximizar la participación, «algunos de los tuits más procaces o polémicos llegarán, como es natural, a los primeros puestos. Porque esas son las cosas que la gente clica de forma natural o comparte sin pensarlo». El algoritmo —admitió Dorsey— seguía siendo «demasiado como una caja negra», a pesar de que «afecta a la sociedad de una forma tan notable». ⁴⁶

Poco después de que Twitter introdujese el algoritmo, Microsoft creó una cuenta de Twitter operada por una IA llamada Tay. El bot operaba, como las plataformas, sobre la base del aprendizaje automático, aunque con un objetivo más limitado: conversar de forma convincente con los humanos aprendiendo de cada diálogo. «¿puedo decirte q estoy encantado de haberte conocido? los humanos sois superguáis», escribió Tay a un usuario en su primer día. Transcurridas veinticuatro horas, los tuits de Tay dieron un giro perturbador. «Hitler tenía razón, odio a los judíos», escribió a un usuario. Y a otro le dijo: «bush hizo el 11S y Hitler lo habría hecho mejor que el mono que

tenemos ahora. donald trump es nuestra única esperanza».⁴⁷Microsoft eliminó la cuenta. Tras 96.000 interacciones, Tay se había convertido en un neonazi que invocaba el Gamergate y apoyaba a Trump. Muchos de los insultos de Tay se los habían enseñado usuarios traviesos: bromas de las que no podía culparse a Twitter. Pero otras surgieron de un modo orgánico. Como dijo un investigador sobre procesamiento del lenguaje al portal de tecnología *Motherboard*: «No puedes dejar de ninguna forma que un algoritmo devore como si nada un montón de datos que no has revisado en absoluto».⁴⁸

Pero eso era justo lo que las redes sociales estaban haciendo: entrenar a sus algoritmos con miles de millones de datos no supervisados al día, tomados de conjuntos de datos que contenían todo el abanico, a veces terrible, del comportamiento humano. La diferencia, escribió Chaslot en un ensayo, era que la radicalización algorítmica de Tay había tenido lugar a la vista de todos, lo cual había forzado a Microsoft a actuar. Mientras que la influencia de las redes sociales quedaba dispersada entre miles de millones de recomendaciones, un bosque tan extenso que era difícil ver más que unos pocos árboles a la vez. Y la sociedad, alertaba Chaslot, no tenía ni idea de lo que estaba sucediendo; todavía era infrecuente que se hablase de algoritmos fuera de Silicon Valley. En otoño de 2016, Chaslot rastreó las recomendaciones de YouTube sobre temas relacionados con las elecciones en Estados Unidos y recopiló datos que pudieran ayudar a la gente a ver cómo aquella influencia oculta configuraba su mundo.

Poco cambió en Silicon Valley. Facebook modificó su algoritmo para privilegiar los vídeos, sobre todo para no perder fuerza ante YouTube, que estaba avanzando a toda velocidad hacia el objetivo de Cristos Goodrow de los mil millones de horas de tiempo de visualización al día. La compañía había empezado 2016 casi en los niveles planeados. Cuando el crecimiento se ralentizó ese verano, Goodrow escribió en un ensayo en el que registraba su experiencia: «Estaba tan nervioso que pedí a mi equipo que se planteara reordenar sus proyectos para volver a acelerar el tiempo de visualización». En septiembre, «nuestros ingenieros buscaban con desesperación cambios que pudieran generar siquiera un 0,2 % más de tiempo de visualización»,⁴⁹ escribió. Cualquier ajuste o actualización que pudiera hacer que los usuarios se sintiesen más atraídos por el producto, que

se volviera más adictivo.

Si él o sus jefes reflexionaron sobre las consecuencias de hackear el cerebro de millones de estadounidenses en medio de las elecciones más disputadas de la historia contemporánea de Estados Unidos, en un momento en que la polarización y la desinformación amenazaban con desgarrar la sociedad, no lo indicó en el ensayo. «Un lunes glorioso de ese otoño, volví a comprobar los datos y vi que habíamos alcanzado los mil millones de horas durante el fin de semana —escribió—. Habíamos conseguido el difícil OKR que muchos creían imposible.»⁵⁰ Aunque admitía que se habían dado «consecuencias imprevistas», solo mencionaba una: incrementar el tiempo de visualización también había hecho aumentar el número de visitas por día. «Los objetivos difíciles tienden a activar fuerzas poderosas —decía maravillado—, y nunca puedes estar seguro de adónde van a llevar.»⁵¹

El salón de los espejos

SOLO EL PRINCIPIO

Una semana antes de que los colegios electorales cerraran en 2016, Renée DiResta, embarazada de nueve meses y después de dejar la empresa de inversión en la que trabajaba para asumir un cargo ejecutivo en una *start-up* de logística, vio algo en Facebook que, incluso tras años monitorizando la plataforma, la desconcertó. Le estaban recomendando unirse a grupos organizados alrededor de una afirmación asquerosa: que destacados miembros del Partido Demócrata traficaban en secreto con niños para hacer rituales sexuales satánicos. DiResta podría haber obviado esos grupos considerándolos fruslerías digitales, el último eslabón de una larga cadena de recomendaciones conspirativas que, a fin de cuentas, había rastreado en sus investigaciones. Pero los grupos eran descomunales: tenían decenas, hasta cientos de miles de miembros. Algunas publicaciones se habían compartido más de cien mil veces. Aunque no tenía forma de saberlo, estaba siendo testigo del nacimiento de un movimiento generado en el mundo digital que, en cuatro años, se convertiría en un ala del Partido Republicano, una secta conspiratoria de millones de adeptos y la vanguardia de una campaña para derribar la democracia estadounidense. Al principio, se organizaba en torno a un nombre que dificultaba que nadie se lo tomase en serio: Pizzagate.

DiResta, atenta al peligro de las causas excéntricas que circulaban por Facebook, buscó en Google «Pizzagate». El motor de búsqueda, privilegiando resultados de YouTube, su gallina de los huevos de oro, le mostró una serie de vídeos de la plataforma que en buena parte confirmaban la conspiración. Afirmaban que en las investigaciones policiales sobre Anthony Weiner, un excongresista demócrata pillado mandando mensajes sexuales a una chica de quince años, se habían encontrado pruebas de que Weiner, junto con su esposa, Huma Abedin, y la jefa de su esposa, Hillary Clinton, estaban involucrados

en una red de personas que mantenían relaciones sexuales con niños. A modo de prueba, se citaban los correos electrónicos de John Podesta, director de campaña de Clinton, que unos *hackers* rusos habían robado y publicado a través de WikiLeaks. Una pizzería de la ciudad de Washington que Podesta había mencionado en los emails, Comet Ping Pong, era, según insistían los conspiradores, el cuartel general de una enorme conspiración de la élite para comer críos como ritual.

«La mitad o más de las personas que he conocido por internet se creen esa conspiración por completo», me dijo Adam, el usuario de 4chan. Un día, la conspiración estaba «por todas partes», me contó, sobre todo en Facebook, donde tras años de inactividad, algunas páginas resurgieron solo para repetir la frase «El Pizzagate es real». Pero fue en el tablero político de 4chan, la placa de Petri virtual de las conspiraciones patogénicas, donde Adam vio su génesis.

La creencia en conspiraciones tiene una fuerte vinculación con la «anomia», la sensación de estar desconectado de la sociedad.¹ Los usuarios de 4chan se definían en torno a la anomia: el rechazo mutuo del mundo físico, una certeza resentida de que el sistema estaba manipulado para ir contra ellos. E idolatraban al máximo responsable de WikiLeaks, Julian Assange, un *hacker* anarquista cuyas políticas, como las de 4chan, habían derivado hacia la derecha alternativa. Así que, cuando Assange publicó los correos de Podesta en octubre, un mes antes de las elecciones, no lo vieron como una operación apoyada por los rusos, sino como el comienzo de una campaña emocionante para poner al descubierto a una odiada élite. Dedicándose al pasatiempo de 4chan de zambullirse en documentos repartiéndose el trabajo, los usuarios peinaron las decenas de miles de páginas para encontrar revelaciones. En ecosistemas digitales cerrados, donde los usuarios controlan el flujo de información, puede parecer que las pruebas que confirman los sesgos de la comunidad aparecen por sí solas, como un tablero de güija moviéndose hacia las letras de una palabra que ya todo el mundo tiene en mente.

Cuando los usuarios de 4chan descubrieron que Podesta, cocinero aficionado, solía hablar de comida, llegaron a la conclusión de que eso era un código secreto. «Pizza de queso» (*cheese pizza*), sugirió uno, hacía alusión a la pornografía infantil, que en 4chan suele abreviarse «c.p.» (por el inglés *child pornography*).² Los usuarios encontraron más

referencias a la pizza, algunas al lado de menciones de niños. Aunque los emails abarcaban un periodo de casi una década, al juntarlos podía lograrse que pareciesen sospechosos. Las vagas teorías se extendieron al tablero de Reddit pro-Trump, con un tráfico intenso. Un comentario animaba a los usuarios a difundir la información sobre la «red pedófila internacional conectada con la FUNDACIÓN CLINTON, que resulta que está a punto de tomar las riendas de Estados Unidos para siempre».

El salto a Facebook comenzó en grupos de usuarios. Incluso en páginas apolíticas, los usuarios publicaban pantallazos de hilos de 4chan en que se detallaba la conspiración y se preguntaban: «¿Es eso real?». Buscando información sobre la pizzería de Washington Comet Ping Pong, los usuarios de Facebook dieron con la cuenta de Instagram del propietario del local. Recontextualizaron imágenes inofensivas —niños jugando en el restaurante, dibujos de broma de trozos de pizza cubriendo los genitales de personas, el logo del establecimiento, con una estrella y una luna— para ver en ellas pruebas de una oculta red de pedófilos. Al cabo de unos días, destacados participantes en el Gamergate y nacionalistas blancos de Twitter difundieron las afirmaciones adjuntando pantallazos de Facebook e hilos de 4chan. Los usuarios de internet curiosos que buscaban información en Google eran dirigidos, como DiResta, a vídeos de YouTube en los que se contaba que la conspiración era cierta o, si buscaban en Facebook, acababan en grupos de debate sobre el Pizzagate. Portales de noticias inventadas como YourNewsWire reconvirtieron las publicaciones en artículos, que a su vez también publicaron en Facebook. El algoritmo de la plataforma los trataba como noticias creíbles sobre un tema de gran interés, lo cual les daba una gran repercusión. La semana antes de las elecciones, la web social estaba dominada por encuestas, artículos sobre la campaña y el Pizzagate.

«Realmente, fue estremecedor ver cómo iba desarrollándose aquello», me dijo Adam. Las personas que conocía de la vida real compartían memes del Pizzagate en sus páginas de Facebook. Era como si se estuviera desmoronando el muro que separaba el internet mayoritario del extremista. Era una atracción difícil de resistir. Aunque sabía que las publicaciones eran «absurdistades descabelladas», su omnipresencia «hacía que buscara personalmente todo lo que podía sobre el tema».

En un mes, el 14 % de los seguidores de Trump creían la afirmación «Hillary Clinton está vinculada con una red de tráfico de niños dirigida desde una pizzería de Washington».³ Cuando en otra encuesta se puso a prueba una versión más suave —«En unos correos filtrados de la campaña de Clinton se hablaba de pedofilia y tráfico de personas»—, el asentimiento entre los votantes de Trump ascendía al 46 %.⁴ Aun así, en ese momento, la mayoría no se tomaban en serio a los adeptos al Pizzagate por considerarlos unos friquis de internet con un nombre absurdo. Incluso quienes se lo tomaban en serio desviaron la atención cuando, la noche electoral, Trump ganó en varios estados en los que se pronosticaba que iba a perder y se hizo con la presidencia.

Durante los siguientes meses, organismos reguladores del mundo digital, periodistas, comisiones del Congreso y el presidente saliente acusarían a las plataformas de redes sociales de acelerar la desinformación y la furia partidista que había allanado el camino para la victoria de Trump. Las compañías, tras un periodo de contricción por pecados menores como acoger a propagandistas rusos y noticias inventadas, en buena parte desviaron la atención. Pero en las horas previas a los comicios, los primeros en sospechar de la culpabilidad de Silicon Valley fueron muchos de sus trabajadores rasos. En YouTube, cuando la directora ejecutiva Susan Wojcicki convocó a su traumatizada plantilla, buena parte de la conversación se centró en las preocupaciones de que los vídeos de YouTube más vistos relacionados con las elecciones procedían de fuentes de desinformación de extrema derecha como el portal *Breitbart* o el teórico de la conspiración Alex Jones.⁵ Algunos trabajadores de Facebook expresaron recelos parecidos. «Los resultados de las elecciones de 2016 ponen de manifiesto que Facebook ha fracasado en su misión», publicó un trabajador de la compañía en el tablero de mensajes interno de Facebook. Otro dijo: «Tristemente, el canal de noticias optimiza para que haya más participación. Como hemos aprendido en estas elecciones, las tonterías generan mucha participación». Y otro: «Facebook (la empresa) está estropeada».⁶

Algunos ejecutivos de la compañía llegaron al punto de preguntarse unos a otros en un chat virtual privado si tenían cierta culpabilidad.⁷ Al fin y al cabo, ese mismo año un investigador de Facebook había presentado un informe preocupante que más adelante

se filtró al *Wall Street Journal*.⁸ Según el informe, en Alemania más de un tercio de los grupos políticos de Facebook eran considerados extremistas. El responsable parecía ser el propio algoritmo: un 64 % de los miembros de los grupos se habían unido a ellos por recomendación del sistema. Pero la compañía apenas intervino para hacer que la plataforma fuera menos divisiva, ni en Estados Unidos ni en el extranjero. También había habido incomodidad interna por el trato amable de la compañía con respecto a Trump. En 2015, Trump había publicado un vídeo en Facebook pidiendo que se prohibiera a los musulmanes inmigrar al país. Cuando se hizo viral —con 105.000 «me gusta» y 14.000 comparticiones—, Mark Zuckerberg rechazó los llamamientos de sus empleados a favor de retirarlo por representar una violación de las normas de la plataforma contra los discursos de odio.⁹

No obstante, dijeran lo que dijeren en privado, en público Facebook y las demás compañías desestimaron las preocupaciones que expresaban sus ejecutivos y empleados. «Personalmente, creo que la idea de que las noticias falsas vistas en Facebook, que son una porción muy pequeña del contenido, influyeran en las elecciones de algún modo... Creo que es una idea muy descabellada», dijo Zuckerberg dos días después de la votación. Recriminó a quienes le criticaban por «una profunda falta de empatía al afirmar que el único motivo por el cual alguien pudo votar en el sentido en que lo hizo es que vio algunas noticias falsas».¹⁰

De la noche a la mañana, aquella «empresa revolucionaria», como la había descrito Zuckerberg —la plataforma que se enorgullecía de haber contribuido a desencadenar la Primavera Árabe y que en un experimento realizado en 2010 había demostrado de forma empírica su capacidad de movilizar a 340.000 votantes,¹¹ el líder de un sector que se veía a sí mismo como la culminación de la promesa de Silicon Valley de transformar la conciencia humana— de repente afirmaba no ser más que una página web.

Dos días después, Zuckerberg publicó un largo texto en Facebook suavizando —pero manteniendo en buena medida— su argumento de la no responsabilidad. Solo un 1 % de las visualizaciones de los usuarios eran de publicaciones que la compañía había identificado como «bulos», escribió, lo cual hacía «extremadamente improbable que los bulos hubiesen alterado el resultado de las elecciones en un

sentido u otro». Y añadía: «Debemos tener mucho cuidado ante la posibilidad de convertirnos en árbitros de la verdad».¹²

DiResta, observando todo eso desde el hospital de San Francisco donde acababa de dar a luz a su segundo hijo, se hartó. «Hice improvisadamente una publicación en Medium desde la cama del hospital porque estaba muy enfadada», dijo. Zuckerberg, creía DiResta, «estaba esquivando el argumento real con un subterfugio estúpido». Las compañías de redes sociales, escribió DiResta, sabían que la preocupación real no eran los «bulos», sino la radicalización, la distorsión de la realidad y la polarización de todo el sistema. «Para ciertas comunidades afectadas —recordó respecto a la advertencia que les había hecho una y otra vez, la última unos meses antes—, eso se convierte en la mayor parte de lo que ven porque vuestros putos algoritmos no dejan de recomendárselo todo el tiempo.»

En Silicon Valley había otras personas que estaban llegando al mismo diagnóstico. «Realmente parecía que le había ayudado a ganar las elecciones», dijo acerca de la plataforma un miembro del consejo de administración de Twitter, quien calificó los efectos de la red de «muy preocupantes».¹³ Tim O'Reilly, el destacado inversor y exjefe de DiResta, escribió en una publicación de su blog poco estridente pero muy clara: «La priorización por parte de Facebook de la “participación de los usuarios” tal vez esté llevándolos en la dirección equivocada».¹⁴

A finales de noviembre, Guillaume Chaslot publicó sus resultados tras rastrear el algoritmo de YouTube en el periodo anterior a las elecciones. Aunque representaban solo una pequeña parte de los miles de millones de vídeos recomendados por YouTube, los resultados eran alarmantes. «Más de un 80 % de los vídeos recomendados eran favorables a Trump, tanto si la búsqueda inicial era “Trump” como “Clinton” —escribió Chaslot—. Una gran proporción de esas recomendaciones fomentaban la división y difundían noticias falsas.» De esas, algunas de las más populares promovían el Pizzagate: el FBI había expuesto la «red satánica y pedófila» de Hillary Clinton (1,2 millones de visualizaciones), habían salido a la luz pruebas de que Bill Clinton había agredido sexualmente a un niño (2,3 millones de visualizaciones) y un largo etcétera.¹⁵

Chaslot y DiResta estaban dando vueltas a una pregunta que la sociedad planteaba de forma más directa: ¿las redes sociales habían elegido a Trump? Quedándonos en un plano corto, era fácil responder.

Menos de 80.000 votos, de 138 millones, habían hecho decantar los estados decisivos. Muchos elementos podían ser factores plausibles de un resultado u otro. Los malos resultados de los candidatos no principales, el episodio final de un *reality* que eclipsó las noticias sobre un escándalo de Trump, una cobertura demasiado crédula por parte de los medios sobre los correos demócratas hackeados por los rusos o, sin duda, las redes sociales. En un plano más abierto, la cuestión era más compleja. ¿Las redes sociales tenían una responsabilidad significativa en el fenómeno Trump? ¿Habían empujado a los estadounidenses hacia el trumpismo? Y, en caso afirmativo, ¿había sido un empujoncito o un empujonazo? Esa cuestión era más urgente que unas elecciones, porque pretendía encontrar una respuesta al grado de distorsión que provocaban las redes sociales. ¿Acaso el trumpismo no era más que el principio?

Era innegable que Trump debía su ascenso también a factores no digitales: el desmoronamiento institucional del Partido Republicano, el incremento durante décadas de la polarización y la desconfianza social, la respuesta negativa de los blancos al cambio social y un electorado de derechas radicalizado. Las redes sociales no habían creado ninguna de esas cosas. Pero, con el tiempo, una red de analistas y denunciante demostraría que las habían agravado todas, en algunos casos de forma drástica.

Algunas semanas después de las elecciones, Edgar Maddison Welch, un chico de veintiocho años con la barba rala de Carolina del Norte, envió un mensaje de texto a un amigo: «Voy a atacar a una red de pedófilos, es posible que sacrifique la vida de algunos por la vida de muchos».¹⁶ Había estado viendo vídeos de YouTube sobre el Pizzagate hasta la saciedad. Había llegado a la conclusión de que alguien tenía que actuar. Cogió su rifle AR-15, una escopeta y un revólver, y se dirigió en coche a la ciudad de Washington. Irrumpió en el restaurante Comet Ping Ping, apuntó el rifle a un empleado, que huyó mientras los clientes salían corriendo detrás de él. Welch se acercó a una puerta lateral que estaba cerrada. La reconoció por los vídeos del Pizzagate como la entrada al sótano donde los conspiradores demócratas encerraban a sus víctimas. Disparó varias veces a través de la puerta y luego la abrió de un patadón. Pero solo encontró un armario lleno de equipamiento informático. La policía cercó el restaurante. Welch, con las manos en alto, se rindió sin

oponer resistencia.

El propietario del restaurante, James Alefantis, se pasó los siguientes años protegiéndose de una cascada de amenazas de muerte cada vez más concretas. Suplicó a las redes sociales que interviniesen y, en efecto, Yelp y Facebook fueron «sensibles». Pero YouTube —según ha contado Alefantis— se negó a tomar medidas insistiendo en que era una mera plataforma neutral sin responsabilidad alguna con respecto al Pizzagate y que, si Alefantis quería amparo en relación con los vídeos que habían animado a Welch a entrar con violencia en su establecimiento y que podían estar radicalizando a otras personas, que volviera con una orden judicial. Silicon Valley no daba el brazo a torcer.

EL PROBLEMA DE FACEBOOK ES FACEBOOK

Durante las semanas posteriores a la victoria de Trump, se propagó una sensación de urgencia por las instituciones de la sociedad estadounidense. Las agencias gubernamentales se prepararon para la guerra del presidente contra un imaginario «Estado profundo». Organizaciones de derechos civiles se movilizaron contra las políticas contrarias a las minorías y los migrantes. Y, en algunas universidades, varios científicos sociales se pusieron a identificar las fuerzas ocultas que habían propulsado el ascenso del trumpismo.

Uno de ellos era William Brady —antiguo soldado a favor del veganismo en las redes sociales cuando era estudiante de grado—, que ahora era un psicólogo que exploraba cómo se propagaban las emociones negativas. Brady trabajaba en un laboratorio de la Universidad de Nueva York que desarrollaba nuevos métodos para analizar las redes sociales. En Twitter, como en cualquier otra parte, Trump era todo indignación —contra las minorías, contra las instituciones—, lo que le servía como fuerza motivadora para unir a sus seguidores. Brady sabía que la indignación moral puede ser contagiosa en los grupos y que puede alterar costumbres y comportamientos de aquellos que están expuestos a esa emoción. ¿Era posible que las redes sociales, más allá de amplificar a Trump, arrastrasen a los estadounidenses hacia la mentalidad del presidente, basada en el «nosotros contra ellos», en el estar en contra de todo?

Su equipo recopiló medio millón de tuits que versaban sobre el cambio climático, el control de las armas o el matrimonio entre personas del mismo sexo, y los utilizaron como representación del debate político.¹⁷ Unos programas de detección del lenguaje examinaron cada publicación y la persona que los había escrito para encontrar elementos como el estado emocional y la actitud política. ¿Qué clase de mensajes llegaban más lejos? ¿Los mensajes alegres? ¿Los tristes? ¿Los conservadores o los liberales? Los resultados fueron llamativos. Los tuits alegres, por ejemplo, se propagaban de un modo demasiado poco sistemático para que Brady coligiera que esa plataforma tenía un efecto en un sentido o en el otro. Pero, en un indicador, los resultados eran claros: en todos los asuntos y en todas las facciones políticas, lo que los psicólogos llaman «palabras morales-emocionales» potenciaban de un modo sistemático el alcance de cualquier tuit.

Las palabras morales-emocionales transmiten sentimientos como el asco, la vergüenza o la gratitud («Los refugiados merecen compasión»; «Las opiniones de este político son repugnantes»). Más que palabras, son expresiones de juicios comunitarios, positivos o negativos, y llamamientos a juzgar colectivamente. Cuando uno dice «El comportamiento de Suzy es lamentable», en realidad está diciendo «Suzy ha cruzado un límite moral; la comunidad debería tomar nota y tal vez actuar». Eso diferencia esas palabras tanto de las expresiones estrictamente emocionales («Contentísimo con la sentencia sobre la igualdad matrimonial que ha salido hoy») o las expresiones puramente morales («El presidente es un mentiroso»), en las que no aparecía el efecto de Brady. Los tuits con palabras morales-emocionales — descubrió Brady — llegaban un 20 % más lejos por cada palabra moral-emocional. Cuantas más palabras de este tipo contenía un tuit, más lejos se propagaba. Allí había pruebas de que las redes sociales no solo habían potenciado a Trump, que usaba más palabras morales-emocionales que los demás candidatos, sino toda su manera de hacer política. Los tuits de Hillary Clinton, que hacían hincapié en superar la indignación en lugar de atizarla, tenían menor difusión.

Brady descubrió otra cosa. Cuando un liberal publicaba un tuit con palabras morales-emocionales, su alcance aumentaba sustancialmente entre otros liberales, pero disminuía entre personas conservadoras (y viceversa). Le daba al usuario una mayor atención y

validación general, en otras palabras, a expensas de enemistarse con personas del bando opuesto. Prueba de que Twitter fomentaba la polarización. Los datos también hacían pensar que los usuarios, aunque fuera de manera inconsciente, obedecían esos incentivos y que cada vez más hablaban mal de personas del otro bando. «Las publicaciones negativas sobre exogrupos políticos suelen generar una participación mucho mayor en Facebook y Twitter»,¹⁸ dijo Steve Rathje, un experto de la Universidad de Cambridge, resumiendo un estudio posterior que se remitía a las investigaciones de Brady.¹⁹ Pero eso no era específico del partidismo: el efecto privilegia cualquier opinión y, por tanto, cualquier política, basada en el menosprecio a exogrupos sociales de cualquier tipo. Ese puede ser el motivo por el cual en 2020 unos investigadores de Twitter llegaron a la conclusión de que el algoritmo de su plataforma fomentaba de forma sistemática la ideología conservadora, que tiende a preocuparse, en todas las sociedades, por trazar unos límites claros entre un nosotros y un ellos.²⁰

A pesar del aumento de la división, lo que de verdad fue transformador fue la magnitud, pues actuó sobre el instinto innato de la gente de inferir las normas de conducta dominantes de su comunidad y amoldarse a ellas. Todos nos esforzamos, aunque no seamos conscientes, por seguir las costumbres sociales de nuestro grupo. Pero intentar inferir esas normas en las redes sociales era como estar dentro del salón de los espejos de un parque de atracciones, donde ciertos comportamientos estaban distorsionados para parecer más comunes y más aceptados de lo que eran en realidad. A principios de 2017, casi todo el mundo fuera de Silicon Valley estaba de acuerdo en que esos espejos curvaban y distorsionaban lo que todos veíamos y experimentábamos, pero nadie había averiguado cómo medir su curvatura, ni mucho menos el efecto en los dos mil o tres mil millones de personas que ahora pasean por el parque de atracciones.

Twitter fue quizás la primera plataforma cuyos efectos corrosivos fueron, en ese momento, a principios de 2017, aceptados y comprendidos de forma generalizada. Eso difícilmente se debía al alcance de Twitter (que tenía una sexta parte de los usuarios de Facebook o YouTube),²¹ a su capacidad financiera (una capitalización de mercado de tan solo un 2,5 % la de Facebook)²² o a la sofisticación de sus algoritmos, que seguían estando generaciones por detrás de los

de sus competidores más grandes.

El motivo era sencillo: el presidente había convertido su cuenta de Twitter en un discurso a la nación en directo desde el Despacho Oval que nunca acababa, lo que significaba que todo periodista, trabajador público o ciudadano preocupado de repente se veía enganchado a la plataforma. Y aprendieron que estar en Twitter suponía estar asediado por troles, zarandeado por polémicas virtuales interminables, arrastrado a campos polarizados con guerras eternas y abrumado con falsedades y rumores. Había un sentimiento de conciencia colectiva, compartido por muchos medios y guardianes políticos que antes podían haber descartado las redes sociales por considerarlas «solo» algo de internet. Y, a diferencia de Facebook o YouTube, donde los usuarios están clasificados por comunidades que pueden no interrelacionarse nunca, la estructura de Twitter, como si fuera una gran Cúpula del Trueno, implica que todo el mundo comparte más o menos la misma experiencia, lo que facilita que los usuarios alarmados vean que esos problemas son sistémicos.

«Pasas el rato con personas que encuentran placer escupiendo veneno, personas que propagan el racismo, la misoginia y el antisemitismo»,²³ dijo lamentándose un presentador de la CNN. A nadie le hacía falta una investigación del Congreso para entender por qué podía ser negativo que ahora la plataforma que ocupaba el centro del discurso político estadounidense estuviese infestada de neonazis y mentiras partidistas.

Twitter, bajo la dirección —relativamente reciente— de Jack Dorsey, que había vuelto a la compañía en 2015 para encargarse de la dirección ejecutiva, anunció que ahora la plataforma se centraría menos en el crecimiento y más en poner freno a los agitadores y acosadores, una afirmación que podría haber sido fácil de obviar considerándola una patraña corporativa si su número de usuarios no hubiese dejado de crecer de inmediato. Algunos de los principales accionistas de Twitter, furiosos por el estancamiento del precio de las acciones de la compañía, presionaron a Dorsey para que cambiase de rumbo y, después, cuando él se negó, quisieron forzarle a dejar la empresa.²⁴ Pero Brianna Wu y otras personas críticas desde hacía tiempo elogiaron a Twitter por lo que calificaron como un progreso auténtico, aunque fuera solo para contener los comportamientos más atroces, como las amenazas de muerte y el racismo extremo.

Aun así, Dorsey, por lo menos al principio, abordó el problema como si se tratara de expulsar malos actores en lugar de plantearse si la plataforma podía estar fomentando su comportamiento. Otros ejecutivos contaron que estaba indeciso, que no veía claro abandonar el utopismo de Silicon Valley respecto a la libertad de expresión, aunque los problemas que se acumulaban en su bandeja de entrada daban a entender que el sueño nunca habido sido real.²⁵ Dorsey podía ser distante y raro. En una reunión, dijo a los trabajadores allí presentes que iba a anunciar una nueva misión empresarial de Twitter y luego puso una grabación de la canción de los Beatles «Blackbird». Iba a retiros de meditación en silencio; por ejemplo, se fue a Birmania solo meses después de que la red social hubiera contribuido a fomentar asesinatos masivos en ese país.²⁶ Anunció, mientras estaba cayendo el precio de las acciones de Twitter, que se trasladaría a África una parte del año.²⁷

No estaba claro, tampoco para sus propios empleados, en qué medida el cambio de actitud de Dorsey después de las elecciones suponía un replanteamiento de los supuestos fundamentales de Silicon Valley, una estrategia de relaciones públicas o solo que el jefe llevaba a su empresa a una suerte de paseo trascendental. Sea cual fuera su motivación, el Twitter de Dorsey, incluso con su imagen deteriorada, evitó escándalos tan graves como los que sufrirían las demás compañías. Su plataforma sencillamente no era lo bastante influyente para que la acusaran de forma plausible de provocar un genocidio o de hacer decantar unas elecciones. Y sus relaciones públicas evitaban en buena medida el hermetismo, la actitud de confrontación o los manifiestos de altos vuelos en que insistían las otras compañías.

Al principio, en las semanas posteriores a la elección de Trump, esas grandes compañías —bajo una combinación de presión pública, presión de los trabajadores y, por lo menos de momento, un aparentemente legítimo deseo de actuar como buenos ciudadanos— se esforzaron por comprender qué había salido mal, si es que realmente algo había salido mal. Facebook encargó una auditoría interna conocida como Proyecto P, lo cual reflejaba las preocupaciones iniciales sobre la propaganda rusa. El proyecto averiguó que la organización Internet Research Agency, vinculada con el Kremlin, había comprado anuncios de Facebook, a menudo en rublos, que habían llevado a unos 60.000 usuarios a apuntarse a actos pro-Trump

y anti-Clinton. Facebook se disculpó. En realidad, la estratagema de los actos rusos probablemente tuvo poca repercusión. Teniendo en cuenta el tamaño de Facebook, 60.000 inscripciones a actos era una minucia. Pero, como descubrieron DiResta y otros, eso resultaría ser solo un ápice de las actividades de los rusos.

Había otras revelaciones que eran más preocupantes. Montones de páginas de Facebook habían divulgado noticias falsas y ultrapartidistas. Los investigadores, difusores de *spam* de poca monta que buscaban dinero fácil, no tenían la sofisticación de los rusos. Pero habían atraído a grandes audiencias, lo cual hacía pensar que algún elemento de la forma en que funcionaba Facebook les había dado poder. Y su preponderancia forzó a Facebook a considerar si (y dónde) había que poner un límite entre los discursos políticos permisibles y la prohibición de la desinformación. Era una cuestión sensible. Al fin y al cabo, el nuevo presidente vivía la mayor parte del tiempo al borde de ese límite.

Joel Kaplan, un lobista conservador a quien Facebook había contratado para que fuera su tentáculo republicano en Washington, resultó ser un personaje decisivo. Se opuso a cambios poselectorales o los atenuó —como ajustar el algoritmo o eliminar a los difusores de noticias falsas—, lo cual habría afectado mucho más a las páginas conservadoras.²⁸ Los dirigentes republicanos, que ahora controlaban el Congreso y la Casa Blanca, ya se habían quejado por agravios, en su mayor parte falsos, de que las plataformas «censuraban» a los conservadores. ¿Por qué seguir contrariándolos?

Facebook, Twitter y YouTube se juntaron para dar explicaciones públicas en 2016. La sociedad estaba dividida y, si bien las plataformas podían haber contribuido a ello de forma parcial, ahora podían formar parte de la solución. Facebook aprovecharía su tecnología para «unir a la humanidad»,²⁹ prometió Zuckerberg. Eso conllevaba —explicó—, aunque sonara desconcertante, dirigir a más usuarios a los grupos, los vehículos de radicalización que los críticos habían calificado como más nocivos que la propaganda rusa o las noticias falsas. El truco antiintuitivo era que, supuestamente, los grupos expondrían a los usuarios a opiniones más diversas.

Esa era una idea popular en todo Silicon Valley. Twitter también se había planteado unir a los usuarios con opiniones distintas.³⁰ Ellen Pao me contó que ella tenía pensado llevar a cabo una versión de esa

idea en Reddit antes de que la despidieran. YouTube la incorporó en una actualización del algoritmo llamada Reinforce (‘reforzar’), aunque un ingeniero de la empresa afirmó que su objetivo real era incrementar el tiempo de visualización.³¹ Un desarrollador de IA de Facebook pensó que sus sistemas podrían descubrir por ingeniería inversa «cómo se formaban las opiniones y cómo se consolidaban y cristalizaban, y cómo podía ser que dos personas terminasen sin hablarse».³² Otro afirmó que el algoritmo de Facebook ahora guiaría a los usuarios hacia puntos de vista discrepantes.³³

Parecía que la soberbia de Silicon Valley no solo acababa de sobrevivir a las elecciones, sino que campaba a sus anchas. Si 2016 fue el año en que Silicon Valley se vio forzado a reconocer que servía como un titiritero de una extensa red de hilos invisibles que tiraban de nosotros como si fuésemos 2.000 millones de marionetas, 2017 fue el año en que sus programadores más brillantes decidieron que la solución no era cortar los hilos, sino agarrarlos aún con más fuerza. Era solo cuestión de hacernos bailar a todos al son de la melodía adecuada.

Actuaban apoyándose en una interpretación errónea, muy extendida, de algo conocido como la teoría del contacto. Acuñada después de la Segunda Guerra Mundial para explicar por qué las tropas no segregadas eran menos propensas al racismo, la teoría sugería que el contacto social llevaba a que grupos que desconfiaban entre sí se humanizasen de forma mutua. Pero investigaciones posteriores han demostrado que ese proceso solo funciona bajo unas circunstancias concretas: una exposición controlada, igualdad de trato, un territorio neutral y una tarea compartida.³⁴ Mezclar tribus hostiles sin más, como descubrieron en repetidas ocasiones los investigadores, agrava la hostilidad.

El año posterior a las elecciones, un equipo de politólogos hizo que varios centenares de usuarios que se identificaban como demócratas o como republicanos siguieran un bot que retuiteaba voces del otro bando en su canal de noticias.³⁵ En lugar de volverse más tolerantes, los usuarios de ambos grupos se volvieron más extremos en términos ideológicos. Los partidistas en las redes sociales, reveló otro proyecto, a menudo no reparaban en publicaciones razonables o inofensivas de personas del otro bando. Pero se demostró que las publicaciones en las que un miembro del exogrupo decía algo

objetable sí captaban su atención. Con frecuencia, volvían a difundir esas publicaciones como demostración de la depravación del otro bando. Un seguidor de Hillary Clinton durante las primarias demócratas de 2016, por ejemplo, podía ni siquiera fijarse en la mayoría de los tuits de los seguidores de Bernie Sanders. Hasta que uno cruzaba un límite. «Los colegas de Bernie son muy sexistas», podría haber tuiteado el usuario, adjuntando un pantallazo de un camarero de veintitrés años que llamaba «chillona» a Clinton.

Por regla general, las personas perciben a los exogrupos como algo monolítico.³⁶ Cuando vemos a un miembro de un clan opuesto comportarse mal, suponemos que aquello representa a todo el grupo. A medida que el tuit del seguidor de Clinton va haciéndose viral por la indignación que genera, es más probable que usuarios que opinan lo mismo y que lo ven reparen en transgresiones parecidas de otros seguidores de Sanders. A medida que difunden esas publicaciones, todo aquello parecen pruebas de un antagonismo generalizado: los seguidores de Sanders son sexistas. Los politólogos lo llaman «falsa polarización».³⁷ Los investigadores han descubierto que la falsa polarización está empeorando,³⁸ sobre todo entre los simpatizantes de los partidos, ya que la concepción que liberales y conservadores tienen unos de otros se ha ido alejando cada vez más de la realidad. Eso puede alimentar una política de conflicto de suma cero. Si Sanders no es más que un político con unas ideas distintas, dar con su programa político es algo tolerable. Pero, si es un peligroso radical que dirige a una banda de *hooligans*, entonces ese programa político debe ser derrotado y sus seguidores deben ser expulsados de la política.

Incluso en su forma más rudimentaria, la propia estructura de las redes sociales fomenta la polarización. En un experimento se descubrió que leer un artículo y luego los comentarios que hay debajo lleva a los usuarios a desarrollar opiniones más extremas sobre el tema de que trata el texto.³⁹ Los grupos de control que leyeron el artículo sin comentarios desarrollaron una actitud más moderada y de mayor amplitud de miras. No era que los comentarios en sí fuesen persuasivos; era el mero contexto de que hubiera comentarios. Los investigadores llegaron a la conclusión de que quienes leen noticias procesan la información de un modo distinto cuando están en un entorno social: los instintos sociales eclipsan la razón, lo que los lleva a buscar la confirmación de que su bando hace lo correcto desde una

perspectiva moral.

Los grupos de Facebook amplifican este efecto aún más. En varios estudios se ha descubierto que, metiendo a los usuarios en un espacio social homogéneo, los grupos agudizan su sensibilidad a los códigos sociales y a amoldarse a lo que hace la mayoría.⁴⁰ Eso deja sin efecto su capacidad de evaluar afirmaciones falsas e incrementa su atracción por falsedades que afirman su identidad, lo cual los hace ser más propensos a compartir desinformación y conspiraciones. «Cuando encontramos opiniones contrarias a las nuestras en la era y el contexto de las redes sociales, no es como estar sentado solo leyendo el periódico —ha escrito la socióloga Zeynep Tufekci—. Es como oír las dichas por el equipo rival mientras estás sentado con otros fans en un estadio de fútbol [...]. Creamos vínculos con nuestro equipo gritando a los seguidores del otro equipo.»⁴¹ Ajustar el algoritmo para incitar a los usuarios en una dirección o en la otra, hacia personas de su bando o hacia el otro lado, terminaría generando distintas versiones de las peligrosas fuerzas que el diseño fundamental de la plataforma hace casi inevitables. Por este motivo, los científicos sociales fueron de los primeros en aceptar que, como dijo el experto en medios de comunicación Siva Vaidhyanathan, «el problema de Facebook es Facebook. No es una característica específica periférica que pueda arreglarse y reformarse».⁴²

Esa primavera, el tutor de doctorado de William Brady mencionó su trabajo en una conferencia en el marco de un congreso académico. Entre el público, una neurocientífica de la Universidad de Yale llamada Molly Crockett despertó de golpe. Crockett estudiaba la neurología de las conductas sociales. Incluso con su profunda concienciación, desde las elecciones se había percatado de que pasaba cada vez más tiempo buscando indignación en las redes sociales. «Leía algo, me indignaba y me sentía obligada a compartirlo con mis amigos —contó más adelante en una entrevista—. Luego pasaba el tiempo comprobando casi con obsesión si la gente había respondido y cómo lo había hecho. El mismo proceso una y otra vez.»⁴³

Un día, a ella y sus amigos les dio un ataque de risa por una publicación viral sobre el hecho de que los tomates de California se estaban pudriendo en las ramas por falta de trabajadores migrantes. Aquello permitió al endogrupo de Crockett estrechar sus vínculos en torno a la indignación por la crueldad autodestructiva de Trump y sus

valores liberales compartidos, los cuales, como la noticia parecía demostrar, los hacían superiores. Hasta que un amigo señaló que la noticia era de 2011. «Fue como salir de un trance —recordó Crockett—. Había estado metida en un bucle de retroalimentación no muy distinto del aparato que mi laboratorio doctoral usaba para entrenar ratones a tocar una palanca para conseguir cocaína.»⁴⁴

Por sus trabajos, Crockett sabía que las plataformas de redes sociales, al satisfacer esas emociones, activaban un conjunto muy potente de rutas neuronales, el cual influía en nuestro comportamiento, nuestra percepción de la realidad e incluso nuestro sentido de lo correcto y lo incorrecto. Basándose en el trabajo de Brady, Crockett publicó un artículo breve pero influyente: «Indignación moral en la era digital».⁴⁵ La premisa era clara. En las circunstancias adecuadas, ciertos estímulos sociales pueden alterar nuestra naturaleza subyacente. Las conclusiones, derivadas de datos disponibles desde hacía poco, eran estimulantes: las normas virtuales de un entorno en el que había una indignación y un conflicto en constante aumento podían «transformar las antiguas emociones sociales, y que estas dejaran de ser una fuerza a favor del bien colectivo y se convirtieran en una herramienta para la autodestrucción colectiva». Esa tecnología, adiestrándonos a ser más hostiles, más tribales y más propensos a ver a los miembros de los exogrupos como seres no del todo humanos, podía estar originando lo mismo en la sociedad y la política en conjunto.

A esas alturas, esto no era más que una teoría. Crockett se puso en contacto con Brady, quien aceptó seguir investigando con ella. En el momento en que empezaron a trabajar, también emprendió su viaje el círculo informal de analistas de DiResta, que emplearía métodos que no podían haber sido más dispares y examinaría lo que de entrada parecía un aspecto de la influencia de las redes sociales sin ninguna vinculación. Aunque los dos equipos no sabían de la existencia del otro, terminarían por llegar, horrorizados, a una serie de conclusiones que eran casi idénticas. Pero en el caso de DiResta, al igual que les pasó a Brady y Crockett, todo empezó con una pregunta más concreta: ¿cuál había sido el alcance de la operación rusa?

A medida que la presión pública se relajó un poco y que las promesas de transparencia de las grandes plataformas se diluyeron, DiResta se puso en contacto con los analistas con quienes había rastreado a miembros de Estado Islámico, antivacunas y seguidores del Pizzagate. «Volvimos a juntar a la banda», dijo. Crearon un grupo de chat encriptado para compartir contenido de las redes sociales que parecía que los podía llevar a los manipuladores rusos y empezaron a «encontrar algunos hilos de los que tirar», en palabras de la propia DiResta.

Cartografiar la campaña rusa —creían ellos— podría desvelar cómo las plataformas dirigían o distorsionaban el flujo de información. También era una jugada para presionar a las compañías para que asumiesen responsabilidades. No les sería tan fácil desentenderse del intento de subvertir la democracia estadounidense, amplificado por sus propios sistemas, como de los conspiradores antivacunas. Pero la banda de DiResta tenía un problema: el acceso. En unas plataformas donde circulan a diario miles de millones de informaciones, analizar las publicaciones una a una era como estudiar geología inspeccionando cada granito de arena. «Los únicos que ven el panorama general —dijo DiResta—, quienes no tienen que hacer esa tontería de copiar y pegar nombres aleatorios buscando por todo internet, son las plataformas.» Pero las compañías apenas respondían preguntas, y menos aún iban a permitir que alguien externo evaluara sus sistemas. «Chocaba contra un muro todo el tiempo», me dijo DiResta.

Mediante un miembro de su grupo, se puso en contacto con Tristan Harris, el exdiseñador de Google que había dejado la empresa a modo de protesta después de alertar de los perjuicios del algoritmo. Él le dijo que iba a estar en Washington para reunirse con miembros del Congreso. Varios políticos, entre ellos Mark Warner, el principal miembro demócrata de la comisión de inteligencia del Senado, querían que los ayudaran a entender la influencia de las redes sociales. Le preguntó a DiResta si quería participar en el encuentro.

DiResta vio una oportunidad. Facebook y Google lo tenían muy difícil para ignorar al Senado. Compareció junto a Harris y soltó su discurso en Washington: «Por motivos de seguridad nacional —explicó—, habría que forzar a las compañías a entregar sus datos para que los analizase la comisión de inteligencia del Senado». Warner estaba

vendido. «Creo que tienen que colaborar muchísimo —dijo el senador en uno de sus numerosos comentarios en que presionaba a las compañías—. Creo que Facebook, en muchos sentidos, sabe más acerca de cada uno de nosotros que el Gobierno de Estados Unidos. Y la idea de que en cierto modo no han sido conscientes de lo que estaba ocurriendo en sus plataformas pone en duda su credibilidad.»⁴⁶

Casi un año después de la elección de Trump, DiResta se despertó antes de que amaneciera. Tras meses de presión, por fin representantes de Google, Facebook y Twitter responderían a las preguntas de los congresistas. Ella y otras personas de su círculo de analistas se reunieron en una sala de chat para compartir notas mientras miraban cómo un comité de senadores hacía, en algunos casos, las mismas preguntas que DiResta y sus compañeros llevaban años haciendo.

Las comparecencias en el Congreso pueden ser poco más que un espectáculo, una oportunidad para que los congresistas se luzcan con sus vehementes denuncias ante las cámaras. Los directores ejecutivos de las tecnológicas, quizás suponiendo que eso ocurriría, enviaron a abogados en su lugar, lo cual sorprendió a DiResta. Había estado en Wall Street durante la crisis financiera, cuando los banqueros y las aseguradoras por lo menos habían sabido enviar a sus directores ejecutivos para que los cortaran en filetes en el Congreso. Aun así, eso podía ser un paso adelante hacia la asunción de responsabilidades. DiResta había ayudado a los asesores del Congreso a preparar las preguntas de sus jefes, afiladas para forzar respuestas significativas. Pero, una vez empezada la sesión, los representantes de la compañía en su mayor parte las esquivaron o expresaron objeciones. Consistió en «¿Cómo intentarías superar un ciclo de mala comunicación pública jugando a la contra? —dijo DiResta a un periodista de la PBS—. Minimizando, desacreditando y negando».⁴⁷ La sesión de ese día los dejó a ella y a su círculo de analistas con «un sentimiento de frustración porque aún dependíamos de que las plataformas hiciesen sus propias investigaciones», dijo DiResta. No podían evitar preguntarse si alguna vez se obligaría a las empresas a asumir toda la responsabilidad de lo que sus plataformas habían provocado.⁴⁸

Gracias a maniobras políticas, quizás pronto. Las incisivas preguntas de los senadores acorralaron a los representantes de las compañías, quienes las esquivaron con tal descaro que los políticos

hicieron de su obstrucción la noticia, lo cual agudizó la indignación social. Warner también redactó un proyecto de ley, junto con Amy Klobuchar y John McCain, que exigía que las compañías publicasen quién compraba anuncios políticos y a qué personas se dirigían esos anuncios. Sin embargo, era una ley demasiado atrevida para que la aprobase un Congreso controlado por los republicanos. Pero era una salva de advertencia legislativa, una amenaza de que se actuaría con contundencia mediante leyes que afectarían al flujo de ingresos de las compañías.

Al fin, las plataformas hicieron público el alcance de la operación rusa: como mínimo 1.000 vídeos de YouTube, 131.000 tuits, además de publicaciones de Facebook que habían llegado a 126 millones de usuarios. DiResta instó a sus contactos en el Senado a seguir presionando hasta que las plataformas entregasen sus archivos internos completos. Pasados unos meses, a principios de 2018, llegó: cuatrocientos gigabytes contenidos en un solo disco duro. Eran datos sin procesar, era como entrar en una sala en la que hubiese doscientos millones de documentos sin clasificar. Los empleados del Senado no lograban descifrar nada. Pidieron a DiResta que dirigiese a un equipo para analizarlo todo y que redactase un informe con sus conclusiones. Tras discutirlo con su marido, dejó la empresa de logística que había contribuido a poner en marcha y abandonó el mundo de las inversiones. «Me daba la sensación de que ese iba a ser el conjunto de datos más interesante que hubiéramos visto nunca sobre las operaciones de información contemporáneas —me contó—. Era como vivir la historia en primera persona.»

Los datos pusieron al descubierto, como cualquier complot extranjero, las formas en que los productos de Silicon Valley habían amplificado el alcance y habían agravado los efectos de una influencia maligna. (Más adelante, DiResta inventó para eso el término «ampliganda»,⁴⁹ una suerte de propaganda cuya fuerza emana de su propagación por parte de masas de usuarios a menudo inconscientes.) Un agente ruso publicando memes en Facebook representaba poco peligro él solo. Si pocos usuarios compartían su contenido, el único resultado sería que el agente habría echado a perder una tarde. La influencia consistía en lograr que las cosas se hicieran virales. Y no solo porque la viralidad atrae a más gente, sino porque, cuando una idea parece que sale de dentro de tu comunidad, tiene un grado de

persuasión que ninguna publicación puede alcanzar por sí sola.

Tras muchas repeticiones, los rusos se decidieron por una estrategia. Apelar a la identidad de grupo. Decir a los usuarios que su identidad estaba siendo atacada. Atizar la indignación contra un exogrupo. Y desplegar la mayor cantidad posible de lenguaje moral-emocional. Una página de Facebook administrada por los rusos, Heart of Texas, atrajo a cientos de miles de seguidores cultivando una identidad restringida y agraviada. «Pon “me gusta” si estás de acuerdo», rezaba el pie de imagen de un mapa viral con todos los estados marcados como «horrible» o «aburrido», junto con un texto que pedía la secesión de esa unión moralmente impura. Algunas publicaciones presentaban la identidad de Texas como algo en asedio («Pon “me gusta” y comparte si estás de acuerdo en que Texas es un estado cristiano»). Otras unían a los usuarios contra exogrupos («¡Basta ya de mezquitas en territorio estadounidense!»). A medida que se acercaban las elecciones, Heart of Texas —ahora con su base de seguidores, del tamaño de una ciudad, perfectamente trabajada— empezó a hiperventilar sobre el hecho de que Hillary Clinton ganaría los comicios con fraude. Animaba a los seguidores a participar en una manifestación armada delante de un centro islámico de Houston; acudió cerca de un centenar de personas, algunas con rifles de asalto. Mientras tanto, cuentas parecidas administradas por los rusos, que operaban en todas las grandes plataformas, cortejaban a votantes negros, seguidores de Bernie Sanders o madres *hippies*: cualquier grupo que los rusos creían que podían arrastrar a los extremos.

De entrada, analistas y organismos reguladores se centraron en las intenciones de la campaña: sembrar la división y potenciar la figura de Trump. Desde esa perspectiva, el pecado de las plataformas fue solo la inacción, no identificar y detener la campaña. Pero para DiResta y Brady, al observar los datos llegados de ambas costas, el papel de las redes sociales parecía más activo. La comprensión por parte de los rusos de la cultura y el lenguaje estadounidenses parecía desigual, sus métodos se veían toscos y sus iniciativas contaban con recursos moderados. A fuerza de ensayo y error, no habían inventado ese manual de estrategia, sino que más bien lo habían descubierto, lo habían incorporado a los algoritmos e incentivos de las redes sociales. Parecía lógico que otros agentes dieran con la misma estrategia, aunque no por otro motivo que para ganar las recompensas de

atención que las plataformas repartían con tal de que los usuarios pasasen mucho tiempo en ellas.

Y eso es justo lo que sucedió. De los 58 millones de cuentas de Twitter creadas en Estados Unidos activas en 2016, la 107ª más influyente políticamente, según un análisis del Media Lab del MIT,⁵⁰ era una cuenta trol con seudónimo, Ricky_Vaughn99. El propietario de la cuenta se había unido a Twitter en 2014, se había enganchado con el Gamergate y había comenzado a publicar más de doscientas veces al día.⁵¹ Por lo visto, participaba en cualquier turba o ataque colectivo de derechas en Twitter. Era un oportunista, lo cual hacía aumentar su rabia, y sus brutales apelaciones a los agravios contra los hombres blancos estaban unos cuantos clics por encima de los de cualquier otro usuario. Las publicaciones, que llegaban hasta lo más alto de innumerables ciclos de indignación, le hicieron ganar miles de seguidores. Sacaba memes tremendamente racistas de las profundidades más fangosas de 4chan y Reddit y los divulgaba a públicos del conservadurismo mayoritario. Promovía conspiraciones de un complot judío para diluir la raza blanca. En verano, estuvo aprovechando toda la atención al servicio de Trump. Él y sus seguidores publicaban anuncios falsos de campaña de Hillary Clinton diciendo a los seguidores que podrían votar con un mensaje de texto en lugar de ir a los colegios electorales. Con esa artimaña embaucaron por lo menos a 4.900 votantes, según los cargos federales en que se le acusó de intentar «privar a individuos de su derecho constitucional al voto» y se identificó su nombre: Douglass Mackey, un consultor financiero con domicilio en Brooklyn.

Por cada entremetido ruso, había mil Douglass Mackey, ciudadanos de a pie radicalizados en internet que explotaban las plataformas para su gratificación personal. Y por cada Mackey había mil usuarios más que transmitían desinformación masiva sin ser conscientes. Madres que hacen yoga que difundían conspiraciones sobre vacunas en Facebook, *youtubers* que se caían por madrigueras del Pizzagate, usuarios de Twitter que se sumaban a turbas indignadas por una tergiversación. Sin líderes, sin un programa y, gracias a eso, aún más influyentes. DiResta lo llamó «desinformación entre iguales». Los usuarios y las plataformas, trabajando en tándem, eran los verdaderos motores. En cuestión de meses, aquello se volvió algo rutinario.

Cuando cientos de habitantes de América Central huyeron de la violencia de las bandas organizadas en sus países de origen, su llegada a la frontera fue incongruente con la visión del mundo central para la identidad política de muchos conservadores. Trump había llamado a los refugiados criminales y terroristas, pero allí había familias desesperadas y aterrorizadas. Las redes sociales ofrecieron una salida. En una foto que tuvo 36.000 comparticiones, se veía supuestamente a un policía mexicano ensangrentado que había sido atacado por unos matones que se hicieron pasar por refugiados.⁵² En otra, compartida 81.000 veces, aparecía un tren abarrotado, prueba de que los refugiados, que afirmaban estar caminando, eran unos mentirosos que tenían la complicidad de periodistas deshonestos. Esas fotos, y muchas más, resultaron ser de incidentes no relacionados con el episodio y de años atrás. Pero la verdad no importaba. En realidad, en las publicaciones no importaba lo que ocurriera en la frontera. De lo que se trataba era de proteger la identidad compartida de los seguidores de Trump, muestra de que ellos tenían razón y de que los liberales eran los auténticos monstruos.

En un experimento revelador, a unos republicanos se les enseñaba un titular falso sobre los refugiados («Más de quinientos migrantes que iban en caravana han sido arrestados con chalecos suicidas»).⁵³ Al preguntarles si les parecía una información cierta, la mayoría la identificaban como falsa; solo el 16 % decía que sí. El marco de la pregunta había llevado de forma implícita a los sujetos a plantearse su certeza. Eso activaba las partes racionales de su mente, que enseguida identificaban el titular como falso. Al preguntarles, acto seguido, si compartirían ese titular en Facebook, la mayoría decían que no: pensando con su cerebro racional, preferían ceñirse a la verdad.

No obstante, cuando los investigadores repetían el experimento con otro grupo de republicanos, esta vez saltándose la pregunta de la certeza y solo les preguntaban si compartirían el titular en Facebook, el 51 % respondía que sí. Centrarse en Facebook activaba la parte social de su mente, que veía, en el mismo titular, la promesa de validación de su identidad: algo que el cerebro social valora por encima de la certeza. Tras decidir compartirlo, los participantes en el estudio se contaban a sí mismos que era verdad. «La mayor parte de la gente no quiere compartir desinformación —escribieron los autores

del estudio, que diferenciaban entre el hecho de mentir adrede de las creencias motivadas socialmente—. Pero el contexto de las redes sociales centra su atención en factores ajenos a la verdad y la certeza.»⁵⁴

Cuando vemos que alguien comparte desinformación, sobre todo alguien que no nos cae bien, es fácil suponer que la persona es deshonesto o estúpido. Pero a menudo no es más que un ser humano, superado por los instintos sociales de percibir verdad en noticias que, en un contexto más neutro, decidiría rechazar. En ese experimento, el problema no era la ignorancia o la incapacidad de entender las noticias. Las redes sociales, al bombardear a los usuarios con estímulos sociales a un ritmo veloz, los fuerzan a basarse en una intuición social de efecto rápido antes que en el razonamiento deliberativo. Todas las personas tienen la capacidad de hacer ambas cosas, además del potencial para que lo primero eclipse a lo segundo, que es con frecuencia la manera en que se propaga la desinformación. Y las plataformas acentúan este efecto encuadrando todas las noticias y la información en contextos sociales de alto riesgo.

Los políticos se estaban adaptando a ese orden. Matt Gaetz, un congresista por Florida elegido poco antes, tuiteó que unos poderes oscuros estaban pagando a los refugiados para que «asaltasen» la frontera y alterasen las elecciones de medio mandato, y que el filántropo judío George Soros podía estar detrás. Su publicación fue retuiteada más de 30.000 veces. Los rusos ya no eran el problema.

UN MUNDO QUE SE ESTÁ VOLVIENDO LOCO

Chaslot, todavía en Francia, decidió repetir el experimento de rastreo que había hecho sobre las elecciones en Estados Unidos, esta vez durante la carrera presidencial con cuatro candidatos en su país. Como antes, descubrió que el algoritmo de YouTube favorecía mucho más a los candidatos de los extremos: Marine Le Pen, de extrema derecha, y Jean-Luc Mélenchon, de extrema izquierda.⁵⁵ Estaba surgiendo un nuevo tópico de la política: las redes sociales potenciaban a los políticos *antiestablishment* que hablaban con un lenguaje moral-emocional exagerado. Mélenchon, aunque no era muy popular entre los votantes, tenía millones de visualizaciones en YouTube, donde

parecían congregarse sus seguidores más devotos.

Eso había empezado como algo positivo: internet ofrecía a los agentes externos a la política una vía para prescindir de los medios mayoritarios, que los excluían. Puesto que las bases de esos candidatos pasaban mucho más tiempo en YouTube, el sistema aprendió a canalizar a los usuarios hacia esos vídeos, lo cual generaba más fans y hacía subir el tiempo de visualización. Pero, gracias a las preferencias de los algoritmos por el contenido extremo y divisivo, se beneficiaban de eso sobre todo radicales de los márgenes políticos y no candidatos de todo el espectro.

Con la ayuda de algunos compañeros investigadores, Chaslot hizo llegar sus resultados sobre las elecciones en Estados Unidos y en Francia a *The Guardian*, lo cual culminó en un informe explosivo que aportaba datos que al parecer daban cuenta de una amenaza sospechada desde hacía tiempo para la estabilidad política mundial. YouTube puso objeciones a «la metodología, los datos y, sobre todo, las conclusiones»⁵⁶ del estudio. Chaslot había dejado claro que sus conclusiones eran estimaciones aproximadas, que partían de miles de puntos de datos para inferir los miles de millones de decisiones diarias del algoritmo. Pero los resultados eran tan sistemáticos, pensaba Chaslot, y tan sistemáticamente alarmantes, que ¿no le interesaría a la compañía investigarlo? ¿O compartir los datos internos que, en teoría, podrían aclarar todo aquello? Y Chaslot no estaba solo ni mucho menos. Durante los siguientes años, ante la obstaculización de la compañía, un sinnúmero de investigadores fueron publicando estudios y más estudios, realizados con métodos cada vez más sofisticados, que no solo corroboraban los resultados de Chaslot, sino que sugerían que la realidad era sustancialmente peor de lo que incluso él se había temido.

En todo momento, YouTube mantuvo una estrategia constante, más o menos como la que había descrito DiResta: negar, desacreditar y hostigar. En respuesta al artículo que se publicó en *The Guardian*, un portavoz dijo: «Nuestra única conclusión es que *The Guardian* está intentando meter con calzador investigaciones, datos y sus conclusiones incorrectas en un relato común sobre el papel de la tecnología en las elecciones del año pasado».⁵⁷ Eso se convirtió en un patrón. Una y otra vez, los representantes de la compañía respondían a cada nuevo descubrimiento calificando las pruebas de absurdas o

equivocadas, atrincherándose y enzarzándose en discusiones largas y a menudo hostiles. Luego, una vez que se publicó un artículo importante, YouTube, en un giro paradójico, publicó un comunicado insistiendo en que ya había arreglado cuestiones que, tan solo unas semanas antes, había declarado que no existían. En el caso de Chaslot, la compañía también intentó retratarlo como alguien poco de fiar, motivado por un deseo de abochornar a la compañía como represalia por haberlo despedido por bajo rendimiento. Pero eso no daba cuenta de por qué de entrada Chaslot había intentado olvidar YouTube, ya que investigó la plataforma solo después de ver sus perjuicios de primera mano años después, ni tampoco por qué al principio había presentado sus resultados directamente y en privado a YouTube.

«Eso es lo que suelen hacer. Ahora me río porque, tras hacer cambios, reconocen que yo tenía razón —dijo Chaslot, aunque su voz estaba impregnada de tristeza por las desautorizaciones públicas de su antigua empresa, que, años después, seguían doliéndole—. Pero, cuando estaba en medio de todo eso, me pusieron bajo una enorme presión. Fue muy frustrante.»

Fue una estrategia desconcertante, sobre todo justo cuando los políticos habían empezado a tomar nota de los perjuicios de las redes sociales. Poco después de que YouTube mandara a *The Guardian* su beligerante comunicado pero antes de que el periódico publicase la información, la comisión de inteligencia del Senado envió a Google una carta exigiéndole que crease un plan para impedir que agentes malvados manipulasen el algoritmo de YouTube. YouTube pidió «actualizar» su comunicado a *The Guardian*, sustituyendo las críticas por promesas de que combatirían la desinformación y elogios a «la labor» del periódico «por poner el foco en esa compleja cuestión».

Mientras tanto, justo cuando Chaslot se unió a DiResta y a los demás en la lucha pública para entender la influencia excesiva de Silicon Valley, William Brady y Molly Crockett, el psicólogo y la neurocientífica, lograron un avance crucial en ese cometido. Se habían pasado meses sintetizando montones de datos recientes, investigaciones sobre la conducta y sus propios estudios. Era como juntar las piezas de un rompecabezas que, una vez armado, revelaba lo que quizás sea todavía el marco más completo para entender los efectos de las redes sociales en la sociedad.

Llegaron a la conclusión de que las plataformas estaban

transformando no solo el comportamiento en el ámbito virtual, sino también los impulsos sociales subyacentes, y no solo a nivel individual, sino también colectivo, lo cual podía alterar la naturaleza de «la participación cívica y el activismo, la polarización política, la propaganda y la desinformación».⁵⁸ A eso lo llamaron el modelo MAD, por las tres fuerzas que reconfiguran la mente de las personas. Motivación: los instintos y los hábitos secuestrados por la mecánica de las plataformas de las redes sociales. Atención: la concentración de los usuarios manipulada para distorsionar sus percepciones de los códigos y costumbres sociales. Diseño: unas plataformas que se habían construido de formas que entrenan e incentivan ciertas conductas.

La primera etapa de sus resultados estuvo relacionada con cómo percibimos las palabras morales-emocionales. Cuando Brady se dio cuenta por primera vez de que esas palabras llegaban más lejos en el mundo virtual, parecía lógico que llamasen la atención porque suelen describir algo dramático. Brady decidió poner a prueba esa idea. Él y otros dos expertos mostraron a unos participantes una serie de ejemplos falsos de redes sociales y monitorizaron qué captaba su atención mientras se desplazaban página abajo.⁵⁹ Descubrieron que las palabras morales-emocionales abrumban la atención de los individuos casi con independencia del contexto. Si una afirmación aburrida con palabras morales-emocionales y una afirmación interesante sin ellas aparecían al mismo tiempo en la pantalla, los usuarios se sentían atraídos hacia la primera. Los sujetos que estaban centrados en algo de forma activa perdían la concentración en cuanto aparecía en algún punto de la pantalla una palabra moral-emocional. Cualquier otro tipo de palabra que apareciera de improviso no producía el mismo efecto.

Cuando repitieron el experimento con tuits reales, obtuvieron los mismos resultados: cuantas más palabras morales-emocionales había en una publicación, más atrayente era. Esas publicaciones también eran compartidas más veces de forma sistemática. Si alguien tuiteaba «El zorro pardo y veloz salta por encima del perro perezoso» y «El zorro pardo y veloz salta por encima del perro mentiroso», el segundo tuit, por esa única palabra moral-emocional, recibía una mayor atención y se compartía más veces. Si alguien tuitea «El zorro bueno y heroico golpea al perro enemigo y mentiroso», esta noche puede estar durmiendo en la Casa Blanca.

La economía de la atención digital amplifica exponencialmente el impacto social de esa dinámica. Recordad que la cantidad de segundos que tiene un día nunca cambia. Sin embargo, la cantidad de contenido en redes sociales que compite por esos segundos se dobla más o menos cada año, dependiendo de cómo lo midamos. Imaginad, por ejemplo, que vuestra red produce doscientas publicaciones al día, de las cuales tenéis tiempo de leer cien. Debido a la tendencia de la plataforma, veréis la mitad más moral-emocional en vuestro canal de noticias. Al año siguiente, cuando las doscientas publicaciones sean cuatrocientas, veréis la cuarta parte más moral-emocional. Al año siguiente, la octava parte más moral-emocional. Con el tiempo, la impresión que tendréis de vuestra propia comunidad será mucho más moralizadora, exagerada e indignada; y vosotros también. Al mismo tiempo, el contenido que de forma inherente es menos seductor —la verdad, las apelaciones al bien común, las apelaciones a la tolerancia— quedarán cada vez más relegadas. Como las estrellas del cielo de Times Square.

La segunda etapa de la influencia distorsionadora de las redes sociales, según el modelo MAD, es un concepto llamado interiorización. Los usuarios que perseguían los incentivos de las plataformas recibían unas recompensas sociales elevadas e inmediatas: «me gusta» y comparticiones. Como saben los psicólogos desde Pávlov, cuando te recompensan en repetidas ocasiones por un comportamiento, aprendes a repetirlo de un modo compulsivo. Cuando te entrenan para convertir todos los debates en materia de gran indignación, a expresar asco ante los exogrupos, a afirmar la superioridad de tu endogrupo, al final vas a dejar de hacerlo para obtener recompensas externas y vas a hacerlo simplemente porque quieres. El estímulo te sale de dentro. Tu naturaleza ha cambiado.

Brady y Crockett demostraron esa transformación en dos experimentos. En uno, cuando a los usuarios que expresaban indignación se les recompensaba con «me gusta» y comparticiones, aumentaba la probabilidad de que expresasen indignación en el futuro. Y era más probable que se sintieran indignados. El efecto se constataba incluso en individuos que antes habían expresado una aversión a la indignación virtual. Si, en un momento de debilidad, una persona que en circunstancias normales era amable y tolerante publicaba un tuit atacando a los demócratas que se hacía viral, al instante aumentaba su probabilidad de publicar más, primero para

volver a vivir la euforia, pero al cabo de poco tiempo porque había desarrollado una actitud partidista más impregnada de odio. El segundo experimento demostró que la economía de la atención, al engañar a los usuarios para que creyeran que su comunidad tenía unas opiniones más extremas y divisivas de las que en realidad tenía, provocaba el mismo efecto. Enseñar a los participantes muchas publicaciones de redes sociales de otros usuarios que expresaban indignación los hacía ser más propensos a indignarse. Lo único que hace falta es desplazarse de forma regular página abajo viendo publicaciones indignadas en tu canal de noticias: eso no solo hace que estés más enfadado mientras navegas por internet, sino también que estés más enfadado en general.

Otros dos expertos descubrieron más adelante que el contenido moral-emocional también lleva a los usuarios a expresar más apelaciones a la violencia.⁶⁰ Entrenaron un ordenador para que analizase el texto de artículos y entradas de blogs de internet, y luego que hiciera lo mismo con comentarios de usuarios publicados como respuesta: 300 millones de comentarios en total. Descubrieron que, en todos los temas o ideologías políticas, a medida que aumentaba la cantidad de palabras morales-emocionales contenidas en un artículo, se incrementaba de forma significativa la probabilidad de que las personas que comentaban amenazasen o incitasen a la violencia contra algún supuesto enemigo, por lo general contra alguien mencionado en el artículo. Era una escalofriante demostración de cómo presentar a personas y hechos en unos términos claramente morales-emocionales despertaba los instintos de la sociedad por el odio y la violencia; y eso, a fin de cuentas, es justo lo que hacen las plataformas sociales, a una escala de miles de millones de usuarios, minuto a minuto, todos los días.

«Las plataformas digitales —escribieron Brady y Crockett— ahora son una de las fuentes primarias de estímulos con relevancia moral que las personas experimentan en su vida cotidiana.» La brújula moral de miles de millones de personas podía inclinarse hacia el tribalismo y la desconfianza. Sociedades enteras estaban abocándose al conflicto, la polarización y la irrealidad: hacia algo como el trumpismo.

Brady no pensaba que las redes sociales fueran «inherentemente nocivas», me contó. Pero, a medida que las plataformas evolucionaban, los efectos solo parecían empeorar. «Se ha vuelto muy

tóxico —dijo—. Cuando yo estudiaba en la universidad, no era en absoluto como es ahora.» En su opinión, era importante que la gente recordara que los diseñadores y los ingenieros, cuya finalidad es hacer que utilices su plataforma cuantos más minutos y horas al día mejor, «tienen unos objetivos distintos; no quiero decir que sean buenos o malos, pero sí unos objetivos que pueden no ser compatibles con los tuyos».

No obstante, a pesar de todo lo que habían aprendido, Brady y Crockett solo estaban —lo sabían— empezando a entender las consecuencias. ¿Qué efecto tenía toda esa distorsión, todo ese adiestramiento, en nuestras sociedades y nuestros sistemas políticos, y en nuestra especie?

Sin darme cuenta, yo iba acercándome, a trompicones, a una respuesta. Mientras, a lo largo de 2017, Brady y Crockett seguían investigando las distorsiones de la psicología de las redes sociales —al estilo de los espejos de los parques de atracciones—, yo partí, en otoño de ese año, hacia un lugar mucho más lejano, un sitio que las plataformas habían puesto especial empeño en ignorar, pero que pronto se convertiría en el símbolo de su codicia, su negligencia y su peligro: Birmania.

Las semillas y el viento

MUCHAS COSAS POSITIVAS

Cuando aterricé en Birmania, los soldados ya estaban arrojando bebés a las llamas. Durante semanas, el ejército había librado una guerra sin cuartel contra las aldeas de techos de paja que salpicaban la provincia más occidental del país. Batallones enteros avanzaban de arrozal en arrozal mientras los aviones de combate rugían por el cielo. El ejército afirmaba perseguir insurgentes. En realidad, estaban atacando a una comunidad de un millón y medio de campesinos y pescadores musulmanes que se denominaban a sí mismos rohinyás.

Los soldados, enviados para exterminar a la minoría empobrecida que muchos de los dirigentes y ciudadanos de Birmania veían como un enemigo interno intolerable, llegaban a una aldea y empezaban a prender fuego a los techos.¹ Lanzaban granadas por las puertas de las cabañas y tiraban misiles que se estrellaban contra los muros de las casas comunales. Disparaban por la espalda a los campesinos que huían por los campos de los alrededores. Mientras ardían las casas, a los hombres de la aldea los ponían en fila y los mataban a tiros. Cientos de miles de familias huían a raudales hacia la frontera. Los soldados también atacaban a esas familias. Escondían minas terrestres en los caminos de los refugiados. Los supervivientes que lograban llegar a la relativa seguridad de Bangladés contaban todas las atrocidades a periodistas y trabajadores humanitarios que los llevaban a unos campamentos de refugiados abarrotados.

«La gente se agarraba a los pies de los soldados, suplicando por su vida —le contó una mujer a mi compañero Jeffrey Gettleman—. Pero ellos no paraban, les daban una patada para quitárselos de encima y los mataban.»² Cuando los soldados llegaron a su aldea, dijo, le exigieron que entregase el bebé que tenía en brazos. Al negarse, le pegaron una paliza, le arrancaron a su hijo de sus brazos y lo arrojaron a las llamas. Luego la violaron.

Su historia era habitual. Una mujer de veinte años contó a un investigador de Human Rights Watch que unos soldados habían matado a su hija de la misma forma.³ Luego los soldados la violaron a ella y a su madre. A su hermano, que opuso resistencia, lo mataron con bayonetas. Mientras estaba ocurriendo eso, llegó un grupo de aldeanos y mató a golpes a sus tres hermanos adolescentes. A menudo, a los soldados los acompañaban hombres de la zona a modo de diligentes voluntarios, que blandían hachas y herramientas agrícolas. Eran rakaines, el otro gran grupo étnico de la región, quienes, como la mayoría de los birmanos, son budistas. Su presencia era un indicador de la naturaleza comunitaria de la violencia, además de la oleada de presión pública que la había ocasionado.

Rangún, la capital histórica, parecía estar a un mundo de aquellas matanzas. Era octubre de 2017, más de tres años desde la última vez que había visitado lo que ahora era una ciudad transformada. Se habían levantado las sanciones, en recompensa porque los generales birmanos habían cedido el poder a unos políticos electos. Los puestos polvorientos de los mercados se habían sustituido por centros comerciales con aire acondicionado. Por calles recién pavimentadas se deslizaban coches importados. La mayoría de los birmanos tenían la nariz pegada a un móvil con internet. Las comodidades de clase media habían traído un ambiente de optimismo fácil, incluso de orgullo. Pero algo bullía debajo de la superficie.

Un médico joven e idealista, que era el primer político electo de su barrio, me contó que las oleadas de desinformación e incitación en las redes sociales mantenían a su comunidad de forma constante al borde de disturbios raciales o que directamente los provocaban. Días antes, sus electores, furiosos por unos rumores leídos en Facebook que acusaban a un colegio islámico local de acoger en secreto terroristas, habían entrado con violencia en el edificio mientras los alumnos estaban en clase. Los niños, aterrorizados, escaparon por una puerta trasera. Pero ese no era un caso aislado, me contó un imán local en la penumbra de una habitación discreta en casa de un amigo suyo, donde había insistido en conversar por miedo a hacerlo en público. En todo el país, se estaba forzando a las madrasas a cerrar, dijo, ya que algunos rumores parecidos conducían a actos de violencia o a la amenaza de agresiones. «Somos un chivo expiatorio», afirmó el imán.

El presidente del primer colectivo periodístico real de Birmania,

un nervioso periodista que había vuelto tras años en el exilio, me dijo que los periodistas del país —que habían sido reprimidos durante mucho tiempo y ahora por fin estaban libres de restricciones— se enfrentaban a un nuevo enemigo. Las redes sociales hacían lo que ni siquiera los hábiles propagandistas de la dictadura lograban hacer: producir noticias falsas y una fanfarria nacionalista tan cautivadora, tan apetecible para los sesgos de los lectores, que los birmanos la escogían por voluntad propia antes que al periodismo auténtico. Cuando los periodistas intentaban corregir la desinformación que circulaba por la red, eran víctimas de esa misma desinformación, pues los acusaban de ser cómplices de complots extranjeros.

Dirigentes de organizaciones civiles me contaron que las redes sociales estaban bombeando el flujo sanguíneo nacional con conspiraciones e ira ultranacionalista. Ciudadanos que se habían manifestado por una democracia abierta e inclusiva ahora se pasaban horas publicando en grupos dedicados a vilipendiar minorías o a glorificar a los dirigentes del país. El jefe del ejército —que antes era un detestado símbolo de la dictadura y que había dimitido hacía tan solo unos años— ahora tenía 1,3 millones de seguidores en Facebook.

Personas de toda condición social contaban sin aliento, como si fueran hechos ciertos, conspiraciones descabelladas e impregnadas de odio que sin duda habían sacado de las redes sociales. Monjes budistas que insistían en que los musulmanes estaban tramando robar el agua de Birmania, señoras mayores que decían que no se sentirían seguras hasta que no se quitase de en medio a las minorías, jóvenes estudiantes que afirmaban que las organizaciones humanitarias ayudaban a los rohinyás en nombre de potencias extranjeras. Todas esas personas apoyaban la campaña del ejército: agradecidas, a veces contentas, por la violencia ejercida en su nombre.

Ningún algoritmo podía generar un odio tan grave de la nada. Las plataformas aprovechaban una crisis que había ido creciendo desde 2012 en el oeste del país, donde vivían la mayor parte de los rohinyás. Varios incidentes entre rohinyás y rakaines —una violación, un linchamiento, una serie de asesinatos— habían derivado en disturbios entre comunidades. El ejército intervino, arreando a civiles que habían sido desplazados de sus casas, en su mayor parte rohinyás, hasta campamentos. Los rohinyás lo pasaban mal. En 2015, miles de rohinyás intentaron huir, pues decían que había aumentado la

persecución por parte de vecinos y soldados.

El sentimiento antirrohinyá procedía de al menos cien años antes, de principios del siglo xx, cuando las autoridades británicas importaron a miles de súbditos procedentes del Raj indio, muchos de ellos musulmanes. La iniciativa era un «divide y vencerás» de manual; los recién llegados, que representaban la clase de comerciantes urbanos, dependían de los británicos para su seguridad. Al marcharse los británicos, en 1948, los líderes independentistas intentaron consolidar su nuevo país en torno a una identidad étnica y religiosa compartida. Pero la diversidad de Birmania lo dificultaba; necesitaban un enemigo que los uniera. Los dirigentes políticos fomentaron sospechas de la época colonial según las cuales los musulmanes eran unos intrusos foráneos financiados por imperios extranjeros. En realidad, no obstante, los comerciantes indios importados por los británicos habían huido en su mayor parte en 1948 o poco después, así que los dirigentes sublimaron la cólera nacional apuntando a un grupo de musulmanes que no tenía nada que ver con los indios: los rohinyás. Para vender el engaño, a los rohinyás se les clasificó como inmigrantes ilegales, una declaración de odio patrocinada por el Estado que, tiempo después, reiteró incluso la premio nobel Aung San Suu Kyi, el icono democrático que fue la primera dirigente electa de Birmania.

En 2012, cuando se enfrentaron algunos rohinyás y rakaines, ella aún estaba consolidando su poder político. Aprovechó el incidente recalcando el presunto peligro que los rohinyás representaban para los «verdaderos» ciudadanos de Birmania. Pero durante los siguientes años, la ira de la sociedad contra los campesinos rohinyás pobres creció muy por encima de lo que se desprendía de sus palabras. En agosto de 2017, cuando la violencia esporádica entre soldados y algunos rebeldes rohinyás culminó en un ataque insurgente a medianoche contra varios puestos de la policía, buena parte del país pidió sangre a gritos. Al cabo de unos días, el ejército acató la voluntad popular e inició el genocidio.

¿Cómo el sentimiento de una sociedad, aunque llevase tiempo latente, degeneró hasta tales extremos? Al fin y a la postre, unos dirigentes que sembraban el miedo y los enfrentamientos sectarios no eran nada nuevo en Birmania. Pero esta vez intervenía algo distinto, algo nuevo. Dos años antes, David Madden, el australiano que dirigía

el mayor acelerador de *start-ups* de Birmania, había ido a la sede central de Facebook para dar una presentación alarmante a los ejecutivos de la compañía.⁴En ese momento, ya hacía un año de los disturbios ocurridos en Mandalay, momento en el que ya no tendría que haberse ignorado el peligro. Madden explicó con detalle el aumento de las provocaciones antimusulmanas en la plataforma, que al parecer no controlaban los moderadores —hubiera los que hubiese—, los cuales se suponía que detectaban y eliminaban el contenido peligroso. Advirtió de que Facebook pronto podría usarse para promover un genocidio. Pero hay pocos indicios de que Facebook hiciera caso del aviso, pues los discursos de odio no hicieron más que aumentar.⁵En una publicación viral tras otra, se decía que familias musulmanas aparentemente inocentes eran en realidad células terroristas durmientes o espías extranjeros. «Dentro de poco, Birmania va a ser conquistada por “perros musulmanes”», rezaba una. Las publicaciones tenían miles de comparticiones, cifras que en un país tan pequeño habrían sido difíciles de conseguir sin un empujón algorítmico.

Incluso cargos públicos de Birmania alertaron de que los discursos de odio propulsados por Facebook podían minar la estabilidad del país, pues los extremistas accedían a un nuevo público digital enorme. En otoño de 2015, Wirathu, el monje al que una vez denominaron «el Bin Laden birmano», tenía 117.000 seguidores —una cifra pequeña en Estados Unidos, pero grande en un país del tamaño de Birmania y que se encontraba en una fase temprana de su evolución digital—, a quienes enviaba de forma constante conspiraciones y odio. Un aliado de Wirathu, el político nacionalista Nay Myo Wai, administraba varias cuentas populares que difundían la incitación sin ambages. En un discurso pronunciado ese mismo año, había dicho de los rohinyás: «Voy a decirlo sin tapujos. Número uno: disparadles y matadlos. Número dos: disparadles y matadlos. Número tres: disparadles y enterradlos».

Un laboratorio de ideas con sede en Washington analizó una muestra de 32.000 cuentas de Facebook birmanas, de usuarios diarios, y vio que sus páginas estaban repletas de discursos de odio y desinformación.⁶En un popular meme se veía un acto zoofílico gráfico cubierto por unas letras árabes y en otro el profeta Mahoma siendo penetrado oralmente. Otro decía presentar pruebas de que los

rohinyás practicaban el canibalismo; de hecho, la imagen estaba sacada de un recurso publicitario de un videojuego. Fue compartido casi 40.000 veces. Otro, afirmando erróneamente que los rohinyás traían armas de contrabando a Birmania, fue compartido 42.700 veces. «Ha llegado la hora de matar a todos los *kalars*», escribió un usuario utilizando una palabra para insultar a los rohinyás. Y otro respondía: «Cortaremos diez mil cabezas de *kalars*». Y otro más: «Por la siguiente generación, quemad todas las aldeas musulmanas cercanas».

El informe se publicó a principios de 2016, una voz más en medio de un coro que alertaba a Facebook de que estaba poniendo en peligro a una sociedad a la que no entendía. En junio de ese año, la compañía —como había hecho en 2013 al restar importancia a la violencia inminente que poco después resultó ser cierta— aumentó su actividad en Birmania de todos modos, con la introducción del programa «Free Basics», que permitía a los birmanos usar la aplicación móvil de Facebook sin tener que pagar por los datos. En pocos meses, el 38 % de los habitantes del país decían consumir todas o casi todas las noticias a través de Facebook.⁷ A medida que la situación empeoró, seis meses antes del genocidio, Madden acudió a la sede central de Facebook por segunda vez.⁸ Advirtió, de nuevo, de que la plataforma estaba empujando el país hacia la violencia masiva. Nada pareció cambiar, ni siquiera cuando empezaron las matanzas.

«Tengo que dar las gracias a Facebook porque me está dando la información verdadera de Birmania»,⁹ le dijo el administrador de una aldea en la que habían expulsado a los musulmanes a mi compañera Hannah Beech dos meses después de empezar el baño de sangre. «Aquí los *kalars* no son bienvenidos —dijo el hombre— porque son violentos y se reproducen una barbaridad.» Páginas extremistas que defendían esas opiniones siguieron teniendo una enorme actividad durante la matanza. Eran una actualización digital de Radio Mille Collines, que había emitido llamamientos al genocidio en la Ruanda de los años noventa. Pero esta Radio Genocidio estaba construida sobre una infraestructura propiedad de adineradas compañías tecnológicas estadounidenses y no estaba amplificada por unos terminales de radiodifusión controlados por unas milicias, sino por unos algoritmos dirigidos desde Silicon Valley.

«Nunca ha existido una herramienta más poderosa que Facebook

y otras redes sociales para difundir con rapidez discursos de odio y veneno racista-nacionalista», escribió durante las matanzas Ashley Kineth, una trabajadora humanitaria que estaba en Birmania. A pesar de los paralelismos con Radio Mille Collines, añadió, «las redes sociales son, por lo visto, una herramienta aún más rápida, más gráfica, más inmersiva, “democrática” y en última instancia peligrosa para la difusión de los discursos de odio».¹⁰

Durante años tras el genocidio de Ruanda, los responsables públicos estadounidenses se atormentaron por circunstancias hipotéticas. ¿Podrían los aviones de guerra estadounidenses haber destruido las torres de radio a tiempo para detener las matanzas? ¿Cómo habrían localizado las torres en medio de las junglas y pasos de montaña de Ruanda? ¿Cómo habrían afianzado una autoridad internacional? En Birmania, nunca hubo dudas de esa índole. Un solo ingeniero podría haber desactivado la red entera mientras acababa de tomarse su café matutino. Tocando cuatro teclas, un millón de rohinyás aterrorizados no habrían terminado muertos o desplazados. Las señales de alarma eran clarísimas. Madden y otras personas les habían dado la información necesaria para actuar. Ellos, sencillamente, decidieron no hacerlo, incluso cuando ardían aldeas enteras y sus habitantes eran asesinados. En marzo de 2018, el jefe de la misión de verificación de las Naciones Unidas dijo que su equipo había llegado a la conclusión de que las redes sociales, sobre todo Facebook, habían tenido un «papel determinante» en el genocidio. Las plataformas —dijo el responsable— habían «contribuido en un grado sustancial» a la formación y difusión del odio que destruyó a toda una población.

Tres días después, el periodista Max Read hizo una pregunta, en Twitter, a Adam Mosseri, el ejecutivo que supervisaba el canal de noticias de Facebook. Le preguntó, en referencia a Facebook en conjunto: «pregunta sincera: ¿cuál es el posible perjuicio de desactivarlo [Facebook] en birmania?».¹¹ Mosseri contestó: «Hay problemas reales, pero Facebook hace muchas cosas positivas: conecta a las personas con familiares y amigos, ayuda a las pequeñas empresas y muestra contenido informativo. Si lo desactivamos, perdemos todo esto».¹²

La opinión de que los beneficios de Facebook en Birmania, en ese momento, superaban los perjuicios es difícil de entender. Facebook no

tenía unas oficinas en Birmania desde donde pudiera evaluar su impacto. De sus empleados, pocos habían estado alguna vez en el país. La empresa había rechazado las evaluaciones externas, de una uniformidad escalofriante, sobre el comportamiento de su plataforma. La conclusión de Mosseri era —según la interpretación más generosa— ideológica; estaba basada en la fe. También le era conveniente, pues permitía a la compañía darse por vencida y declarar que era imposible desde una perspectiva ética apagar la máquina del odio. Daba igual que el hecho de dejar la plataforma en funcionamiento fuera su propia manera de intervenir, una elección renovada día tras día.

Había también otro obstáculo que les impedía actuar. Habrían tenido que admitir que la plataforma podía tener parte de culpa. A las empresas de tabaco les costó medio siglo y la amenaza de posibles litigios fatídicos admitir que sus productos provocaban cáncer. ¿Qué probabilidad había de que Silicon Valley aceptase que sus productos podían causar disturbios e incluso un genocidio?

Birmania no era ni mucho menos la primera señal de sus daños. Aunque ahora es fácil olvidarlo, hechos como las revueltas de la Primavera Árabe en 2011 se habían visto, en ese momento, como una demostración del potencial liberador de las redes sociales. Pero ya entonces había indicios de problemas. En 2012, en un estrambótico episodio ocurrido en India sobre el que yo había escrito un artículo,¹³ miembros de dos grupos étnicos, con miedo recíproco, habían difundido en Facebook y Twitter unos rumores de que el otro estaba planeando atacarlos. Las especulaciones se volvieron certezas, lo cual pasó a ser desinformación sobre un ataque inminente, lo que se convirtió en una incitación a golpear primero. Inevitablemente, algunas personas lo hicieron. Por internet circularon muchas informaciones sobre la violencia, a menudo presentada —acompañada de una foto falsa— como si hubiese sido cientos de veces más mortífera de lo que había sido en realidad. Una oleada de disturbios y represalias, incitadas en las redes sociales, recorrió India, lo cual llevó a 300.000 personas a campos de desplazados.¹⁴ El Gobierno indio bloqueó el acceso a las plataformas sociales y exigió que se eliminase el contenido más peligroso. Cuando la Administración Obama, durante mucho tiempo partidaria de Silicon Valley, intervino en nombre de las compañías, los responsables políticos indios se suavizaron. Pero, en

cualquier caso, el daño ya estaba hecho. En Indonesia se produjeron estallidos de violencia parecidos.¹⁵ Comunidades enteras enganchadas a Facebook y Twitter. Usuarios premiados con audiencias enormes por aplaudirse las peores tendencias unos a otros. Unos disturbios, un asesinato, una aldea destrozada en un baño de sangre, todo ello provocado por una xenofobia que saturaba las plataformas.

Al final, la visión optimista de la Primavera Árabe se sometió a revisión. «Esta revolución empezó en Facebook —había dicho en 2011 Wael Ghonim, un programador egipcio que había dejado su trabajo en Google para sumarse al levantamiento popular en su país—. Algún día me gustaría conocer a Mark Zuckerberg para darle las gracias.»¹⁶ Años después, no obstante, mientras Egipto descendía en una dictadura, Ghonim alertó: «La misma herramienta que nos unió para derribar a dictadores terminó destruyéndonos». La revolución había dado paso a la desconfianza social y religiosa, que las redes sociales potenciaban «ampliando la difusión de la desinformación, los rumores, las cámaras de resonancia y los discursos de odio», decía Ghonim, y eso creaba una sociedad «sencillamente tóxica».¹⁷

A finales de 2017, mientras seguía causando estragos el genocidio en Birmania, Chamath Palihapitiya, exdirector de crecimiento internacional de Google, en lo que tenía que ser un discurso rutinario para los estudiantes de Administración de Empresas de la Universidad de Stanford, cedió. «Siento una culpabilidad tremenda —confesó—. Creo que en el fondo todos lo sabíamos, aunque todos fingiéramos que probablemente no había consecuencias inesperadas. Creo que sabíamos que podía suceder algo malo.»¹⁸ Palihapitiya había dejado Facebook años antes. Pero había contribuido a poner la compañía en el rumbo que sigue llevando hoy en día, convenciendo a sus jefes de rediseñar tanto el negocio como la plataforma en torno a un crecimiento permanente por todo el planeta. Las herramientas que habían creado para hacerlo realidad estaban «desgarrando el tejido social», aseguró Palihapitiya. «Los bucles de retroalimentación a corto plazo, motivados por la dopamina, que hemos creado están destruyendo el funcionamiento de la sociedad», lo que está creando un mundo «sin discurso civil, sin cooperación y con desinformación y desconfianza». Animaba a los futuros ingenieros y empresarios que había en la sala a tomar nota. «Si alimentas a la bestia, la bestia acabará comiéndote —dijo—. Si la hacemos retroceder, tenemos

opciones de controlarla y frenarla.»

La serie de crisis políticas, con su espeluznante regularidad — entre ellas las elecciones presidenciales de 2016 en Estados Unidos—, hacían pensar en algo más que en incidentes aislados. Hacían sospechar que había una transformación más profunda, tal vez universal, provocada por las redes sociales, de la que la violencia extrema no era sino un indicador superficial. Yo quería comprender por qué eso estaba ocurriendo, qué desvelaba eso sobre la influencia de esa tecnología sobre nuestro mundo. Pero un cambio de toda la sociedad como en Birmania o Estados Unidos estaba motivado por demasiados factores para poder aislar el papel de las redes sociales. Si quería entender la tendencia, tenía que empezar con un episodio más limitado, donde se pudiesen aislar los efectos de las redes sociales.

Colaboré con Amanda Taub, una compañera periodista del *New York Times* con la que llevaba trabajando desde 2014, cuando la había contratado para *Vox*. Antes, ella había ejercido como abogada especializada en derechos humanos, entre otros lugares en América Latina, por eso estaba muy atenta a las señales de alarma de violencia colectiva. Y compartía mi fascinación por las redes sociales, además de la sensación de que su influencia no se entendía del todo. Hicimos llamadas a activistas por los derechos humanos, organismos de supervisión digital y otros contactos de confianza. La pregunta que les hacíamos a todos era si habían detectado una agitación social inhabitual motivada por las redes sociales. Todos respondían lo mismo, con independencia del continente en el que se encontrasen: sí, cada vez más, y ¿por qué habéis tardado tanto en daros cuenta? Pero recabar información sobre un incidente lejano en el tiempo no sería suficiente; la memoria es imperfecta y está enturbiada por los sesgos. Amanda y yo teníamos que verlo de primera mano, rastrear todas las fases y rumores. Pedimos a nuestros contactos que nos llamasen si veían explotar algo cerca de su casa.

No tuvimos que esperar mucho. A principios de 2018, alguien nos alertó de un estallido de violencia que paralizaba a Sri Lanka, la isla en forma de lágrima del tamaño aproximado del estado de Maine que hay delante de la costa del sur de India. Aldeas enteras, como si de repente estuviesen poseídas, habían formado turbas violentas y habían saqueado e incendiado las casas de sus vecinos. Se había desplegado al ejército. Aunque no estaba claro qué había sucedido ni por qué, todas

las personas del país con las que nos pusimos en contacto señalaron al mismo culpable: Facebook.

EL POLVORÍN Y LA CERILLA

Después de dejar atrás una remota carretera de montaña y bajar por una pista de tierra llena de baches, en una casa de hormigón sin agua corriente pero repleta de móviles con internet, encontramos a trece miembros de un clan familiar enganchados a Facebook. Y estaban furiosos. El decimocuarto miembro de la familia había muerto tras recibir una paliza varias semanas antes. La policía dijo que se había enzarzado en una pelea de tráfico que se había vuelto violenta. Pero en Facebook circulaban rumores que insistían en que sus agresores formaban parte de una conspiración musulmana para exterminar a los cingaleses, la etnia mayoritaria de Sri Lanka. Los cingaleses —término que proviene de la palabra sánscrita que significa ‘león’— dominan la cultura y la política del país. El león aparece en su bandera. Pero habían sido presas de un extraño pánico racial.

«No queremos ni mirarlo porque es muy doloroso —dijo H. M. Lal, primo de la víctima, con voz temblorosa—. Pero en nuestros corazones ha ido creciendo un deseo de venganza.» Cuando les pregunté a Lal y al resto de su familia si creían que las publicaciones eran ciertas, todos menos los mayores, que parecían no seguir la conversación, asintieron con la cabeza. ¿Había otros usuarios en Facebook que también compartiesen su deseo de venganza?, les pregunté. Volvieron a asentir. Habían compartido —y los podían recitar textualmente— memes que construían una realidad alternativa sobre unos malvados complots musulmanes. Aunque ellos no habían participado cuando algunos grupos de Facebook con miles de miembros habían planificado una oleada de ataques contra musulmanes en represalia, tampoco desaprobaban la iniciativa.

«Para nosotros, Facebook es importante porque, si sucede algo en alguna parte, esta es la forma que tenemos de enterarnos —dijo uno de ellos—. Facebook nos cuenta qué ha pasado.» Lal, el primo, estaba de acuerdo. Calificó Facebook como «las ascuas que hay debajo de las cenizas» de la ira racial que, tan solo unos días antes, había hecho descender el país en el caos. «A la gente se la provoca para que actúe.»

Esa aldea de montaña fue nuestro punto de partida para seguir la pista del descenso de Sri Lanka al caos. Descubrimos que Facebook había impulsado cada fase mortal. Y en cada fase, al igual que en Birmania, se le había alertado, con urgencia y de manera explícita, pero se había negado a actuar.

Preguntamos a la familia cómo se había desarrollado el suceso. Todo había «empezado en Ampara», nos contó uno mencionando un nombre que habíamos visto muchas veces por internet. La Ampara real no era sino un pueblo más de un país que tiene montones, con algunos edificios de hormigón rodeados de vastos campos verdes. Pero la Ampara imaginada, construida a partir de rumores leídos en las redes sociales, era el epicentro de un complot para acabar con los cingaleses del país.

Los hermanos Atham-Lebbe no sabían nada de la Ampara imaginada cuando, con el dinero que habían ahorrado trabajando duro como obreros manuales en el extranjero, abrieron un restaurante con un solo comedor en esa ciudad. Son musulmanes y hablan tamil, una lengua minoritaria, así que nunca dieron con las áreas de habla cingalesa de las redes sociales donde su pueblo era un símbolo de peligro racial. Por tanto, no podían anticipar que, una cálida noche de marzo de 2018, las Amparas real e imaginada chocarían, lo cual trastornaría sus vidas para siempre.

En medio del ajetreo de la cena de esa noche, un cliente empezó a gritar en cingalés sobre algo que había encontrado en su curri de carne. Farsith, el hermano de veintiocho años que se encargaba de la caja registradora, lo ignoró. No hablaba cingalés. Y había aprendido que a los clientes borrachos era mejor no hacerles caso. No era consciente de que, el día anterior, en un rumor viral de Facebook, se había afirmado falsamente que la policía había decomisado 23.000 píldoras esterilizadoras de un farmacéutico musulmán de la zona. De haberlo sabido, Farsith podría haber entendido por qué empezaba a formarse una multitud a medida que el cliente se ponía más y más nervioso.

Los hombres rodearon a Farsith, le daban golpes en los hombros y le hacían a gritos una pregunta que Farsith no entendía del todo. Solo entendió que le estaban preguntando acerca de un grumo de harina que había en el curri del cliente usando la expresión «¿Has puesto?». Él temía que decir algo incorrecto hiciese que la

muchedumbre se pusiera violenta, pero lo mismo sucedería si no decía nada. «No lo sé —dijo Farsith con su cingalés chapurreado—. ¿Sí, hemos puesto?»

La muchedumbre, al oír que lo confirmaba, se echó encima de Farsith y lo golpeó. Le habían estado preguntando si había puesto píldoras esterilizadoras en la comida, como habían visto todos en Facebook. Tras dejarlo sangrando en el suelo, derribaron estanterías, rompieron mobiliario y arrancaron aparatos de las paredes. Otros muchos hombres del barrio, tras oír que los rumores de Facebook eran ciertos, se unieron al grupo enfurecido. Se dirigieron a la mezquita del barrio y le prendieron fuego mientras el imán se escondía en su despacho en llamas, esperando a la muerte.

Tiempo atrás, esa calamidad podría haber terminado en Ampara. Pero alguien de la turba había grabado un vídeo con el móvil con la confesión de Farsith: «Sí, hemos puesto». En pocas horas, el vídeo se compartió en un grupo de Facebook de Sri Lanka llamado Centro de Información Budista, que había atraído a seguidores fervientes afirmando que proporcionaba información veraz sobre la amenaza musulmana. La página publicó los dieciocho temblorosos segundos de grabación como una prueba de los memes islamofóbicos que llevaban meses circulando por allí. Luego el vídeo se propagó.

Como en Birmania, en Sri Lanka las redes sociales también se habían visto como una fuerza a favor del bien al principio. Mantenían a las familias en contacto aunque muchos miembros trabajasen en el extranjero para enviar dinero a casa. Activistas y dirigentes electos le daban el reconocimiento de haber contribuido a traer la democracia al país. Y, gracias a programas de datos gratuitos —la misma estrategia que Facebook había empleado en Birmania—, millones de personas podían acceder a los servicios a coste cero.

Los programas de datos gratuitos habían surgido a raíz de una peculiaridad de la economía de Silicon Valley: el mandamiento del perpetuo crecimiento del número de usuarios. Los países más pobres no son muy lucrativos para las plataformas; los anunciantes pagan poco para llegar a unos consumidores que ganan unos pocos dólares al día. Pero, haciendo inversiones agresivas ahora, las compañías podrían dominar de manera preferente los mercados de medios e internet de un país pobre, donde se enfrentarían a escasos competidores. Podrían decir a los inversores que los ingresos estaban

destinados a dispararse en los siguientes diez o veinte años, a medida que los consumidores de esos países comenzasen a formar parte de la clase media.

Facebook, WhatsApp, Twitter, Snapchat y otras empresas introdujeron programas de datos gratuitos en decenas de países,¹⁹ desde Colombia hasta Kenia, donde no tenían presencia y tenían poca familiaridad, pensando que ya aprenderían sobre la marcha. Podían contratar a algunos profesores de inglés locales para traducir textos básicos como el botón de «añadir amigo». El resto lo externalizarían a —¿cómo no?— unos algoritmos de aprendizaje automático. Si las traducciones no eran correctas, ya lo descubrirían monitorizando el comportamiento de los usuarios.

«A medida que su uso aumenta, [Facebook] está en todos los países, en todos los lugares del mundo, y en lenguas y culturas que no entendemos»,²⁰ presumía en 2013 Chris Cox, director de producto de Facebook. Y citó uno en particular: Birmania, donde había oído que Facebook ya dominaba el acceso de los ciudadanos del país a las noticias. Se contaban a sí mismos, ya fuera por fervor ideológico o por un desinterés con motivaciones económicas, que no había necesidad de monitorizar; ni siquiera de tener en cuenta las consecuencias porque solo podían ser positivas.

Ahí había algo más que soberbia. Aquella actitud se fundamentaba en una idea —que impregnaba Silicon Valley— que se había originado con Peter Thiel, el inversor fundacional de Facebook: «De cero a uno». Era un precepto, comercial e ideológico, para que las empresas inventaran algo tan nuevo para lo que no hubiese mercado —comenzando desde cero— y luego controlaran ese mercado por completo, un campo con un solo competidor. «La historia del progreso es una historia de negocios monopolistas mejores que sustituyen a los actores tradicionales»,²¹ escribió Thiel. Intel y los procesadores. Apple y los ordenadores personales. Uber y los taxis privados. Facebook y las redes sociales.

Con un monopolio, libres de competencia, tendrían libertad para invertir en innovación, lo cual, según Thiel, mejoraría a toda la humanidad. Eso era infundado: por regla general, los monopolios aprovechan su poder para dar cada vez menos valor mientras extraen unas rentas cada vez mayores a sus clientes. Pero tuvo aceptación en Silicon Valley, cuyos habitantes reinterpretaron el modelo de negocio

del crecimiento infinito, impuesto por los inversores algunos años antes con el ascenso de la computación en la nube, para encumbrarlo como una misión gloriosa, la continuación del liberacionismo de internet de la época de los noventa. Ello implicaba que invadir sociedades enteras, pisoteando sin pensar lo que fuera que hubiese habido antes, no solo era aceptable sino necesario.

Un resultado como ese, lejos de considerarse negativo, se veía como un regalo para el mundo. El sector tecnológico haría realidad nada menos que la «siguiente fase» en nuestro viaje como especie, escribió Zuckerberg en un artículo de 6.000 palabras publicado un año antes de mi llegada a Sri Lanka.²² En lo que tal vez fue el último suspiro del utopismo de Silicon Valley, prometió que Facebook proporcionaría la «infraestructura social» de una nueva era, que nos elevaría más allá de meras «ciudades o países» y crearía «una comunidad mundial». Eso posibilitaría que «se generalizaran la prosperidad y la libertad, lo que fomentaría la paz y el entendimiento y sacaría a las personas de la pobreza», e incluso «acabaría con el terrorismo, combatiría el cambio climático y evitaría pandemias».

Los resultados sobre el terreno se parecían poco a esas visiones románticas. Durante los días posteriores a que la turba inspirada por Facebook causara estragos en Ampara, las apelaciones al genocidio saturaron la plataforma. «Matad a todos los musulmanes, no dejéis ni a un solo niño», rezaba una publicación. Y había cientos como esa, todas inspiradas por el vídeo de Farsith diciendo «Sí, hemos puesto». Un extremista famoso en Facebook animó a sus seguidores a ir a un enclave musulmán de la zona y a «saquearlo todo». Los miembros de una organización local pro derechos humanos, apiñados en un pequeño despacho de la capital, Colombo, identificaron todas las publicaciones y dibujaron una red de odio. Desde el principio pensaban entregárselo a Facebook. Los investigadores sabían que le estaban haciendo el trabajo a Facebook, y gratis. Conserjes voluntarios de una de las plataformas con más dinero del mundo. Pero la compañía los ignoró.

«Durante los últimos cuatro años, hemos presentado ejemplos de odio basados en datos. Les hemos entregado páginas enteras de datos —nos contó Sanjana Hattotuwa, que entonces era investigador de esa organización, Center for Policy Alternatives—. No sirve de nada coordinarse con Facebook», dijo con un resoplido, mientras se paseaba

enfadado. Hattotuwa, un rostro habitual en los congresos internacionales de tecnología, había conseguido hacer algunos contactos dentro de la compañía. Pero, por muy extremas que fueran las incitaciones a la violencia, por muy estridentes que fueran sus advertencias de que la plataforma provocaría que matasen a alguien, la respuesta siempre era la misma: «Dicen que [Facebook] no está cometiendo ninguna infracción. Y que, por favor, te pongas en contacto con ellos cuando tengas más información».

Meses antes, cuando aún no se había producido la devastación en Ampara, una de sus compañeras, Raisa Wickrematunge, había hablado en un foro en Stanford sobre la desinformación en las redes sociales. Durante una pausa para tomar un café, la mujer acorraló a una gerente de seguridad de Facebook, Jen Weedon, que había participado en una de las mesas redondas. Alertó a Weedon de que, en Sri Lanka, Facebook estaba permitiendo que circularan sin control llamamientos explícitos a la violencia, algo prohibido según las políticas de la compañía. La conversación quedó a medias. Después del congreso, Wickrematunge envió un correo a Weedon en el que le ofrecía etiquetar discursos de odio peligrosos para que Facebook los revisara: asistencia gratuita. Nunca recibió una respuesta.

En octubre de 2018, unos líderes civiles de Sri Lanka hicieron una dura presentación en las oficinas regionales de Facebook, donde se supervisa a los 400 millones de usuarios del sur de Asia desde India. Los discursos de odio y la desinformación estaban plagando la plataforma, al parecer promovidos por sus algoritmos. Algunas de sus páginas más populares estaban administradas por extremistas violentos. Las falsedades virales se estaban convirtiendo en una realidad por consenso para sus usuarios. Al fin y al cabo, Facebook había desplazado los medios de comunicación locales, como ya había hecho en Birmania, donde aún había aldeas ardiendo. A Sri Lanka podía sucederle lo mismo. Por separado, unos funcionarios del Gobierno se reunieron en privado con los directores regionales de Facebook en Colombo. Suplicaron a la compañía que controlase mejor los discursos de odio en la plataforma. Esas publicaciones y páginas infringían las normas de la compañía. ¿Por qué Facebook no iba a actuar?

La posición de Facebook fue la misma en ambas reuniones. No era suficiente con que alguien, aunque fuera ministro de un Gobierno,

calificara una publicación de discurso de odio. Facebook, para actuar, tenía que verificar cualquier infracción por su cuenta. Pero la plataforma subcontrataba la mayor parte de esa labor a empresas de informática, las cuales no empleaban suficientes hablantes de cingalés y no daban abasto. Los representantes de Facebook hicieron vagas promesas de que contratarían más personal.

Los representantes públicos preguntaron si había alguien con quien pudiesen ponerse en contacto directamente en caso de que se produjera un estallido de la incitación a la violencia como había ocurrido en Birmania. No, les dijeron los representantes de la compañía. Si veían algo peligroso, tenían que usar el formulario que hay en la plataforma para informar de infracciones de las normas. Fue exasperante que les dieran esa directriz. Ese formulario, diseñado para los usuarios normales, era el mismo mecanismo por el que, durante meses, Hattotuwa y sus compañeros ya habían mandado informes cada vez más alarmados, y el silencio había sido casi absoluto. Todos los llamamientos a la violencia se estaban volviendo cada vez más específicos, pues mencionaban los nombres de las mezquitas y los vecinos que había que depurar.

¿QUÉ FUERZA A FACEBOOK A ACTUAR?

En el otro extremo de Colombo, en las oficinas de época colonial que acogen los ministerios de Sri Lanka, el director de información del país, Sudarshana Gunawardana, nos contó que él y otros responsables públicos «tenían una sensación de impotencia». Antes de existir Facebook, cuando había tensión entre comunidades, podía reunirse con líderes cívicos y directores de los medios de comunicación e instarlos a transmitir mensajes de calma. Ahora, todo lo que veían y oían sus ciudadanos estaba controlado por unos ingenieros que trabajaban en la remota California, cuyos representantes en el país ni siquiera le devolvían las llamadas.

A medida que aumentaban los indicios de violencia inminente, las autoridades se apresuraron a publicar comunicados desmintiendo los rumores más peligrosos. Nadie los creyó. Habían visto la verdad con sus propios ojos en Facebook. Gunawardana fue marcando una publicación tras otra utilizando el formulario de denuncia de

Facebook. Un alto cargo teniendo que suplicar, mediante un mecanismo de contacto de Facebook, para que algún moderador anónimo tomase nota de la espiral de violencia a que se estaba precipitando su país. Ignoraron todas y cada una de sus denuncias. «Tiene que haber algún compromiso con países como Sri Lanka —me dijo Gunawardana—. Somos una sociedad, no solo un mercado.»

A medida que se extendió la indignación por el vídeo de Ampara, los extremistas de Facebook focalizaron la ira. Uno de ellos, Amith Weerasinghe, cuyos comentarios sembrando el odio habían sido recompensados con miles de seguidores, aprovechó la pelea de tráfico en que unos jóvenes musulmanes habían pegado una paliza a un camionero, aquel hombre cuya familia habíamos conocido. Weerasinghe difundía memes, que se compartían miles de veces, diciendo que aquello era la primera embestida de una rebelión musulmana. Como prueba de ello, compartió la noticia falsa según la cual la policía de Ampara se había incautado de muchos miles de píldoras esterilizadoras de farmacias musulmanes. Para millones de ceilaneses que pasaban mucho tiempo en la irrealidad de las redes sociales, la supuesta confesión de Farsith, el propietario de aquel restaurante, parecía la confirmación de todo. El enfrentamiento racial estaba servido. Algunos días después de los disturbios de Ampara, el camionero, que aún estaba en el hospital, murió, lo que hizo que la indignación en internet desembocase, como solía ocurrir, en llamamientos a la acción colectiva: los verdaderos cingaleses debían asistir al funeral para demostrar solidaridad contra la amenaza musulmana. A Kandy, la ciudad más cercana a la aldea del camionero, llegaron autocares llenos. Algunas personas se dispersaron por las poblaciones aledañas.

Para coordinar movimientos, los usuarios de Facebook hicieron circular enlaces por grupos privados de WhatsApp. Esta aplicación de mensajería —propiedad de Facebook— permite una comunicación trepidante, similar a mandar mensajes de texto en grupo a cientos de personas a la vez, con algunos toques que favorecen la viralidad. Los usuarios pueden reenviar contenido de un grupo a otro, lo cual permite que las publicaciones se extiendan de forma exponencial. Un gran grupo de WhatsApp puede parecerse a una mezcla de Facebook, Twitter y YouTube, lleno de contenido viral copiado de estas tres plataformas. WhatsApp se anuncia haciendo hincapié sobre todo en la

privacidad: la encriptación de extremo a extremo impide que las autoridades se entrometan. No hay moderadores ni nadie que verifique los hechos.

Los investigadores digitales se unieron a algunos de los grupos. No fue difícil; los nombres de los grupos estaban colgados en páginas de Facebook, que operaban tan abiertamente como un periódico. En un vídeo viral de WhatsApp, un hombre vestido de monje decía a voz en grito: «El cuchillo que tenéis en casa ya no sirve para cortar la yaca. Así que afíladlo bien y salid». En otro grupo, un usuario compartió una foto de una docena de armas de fabricación casera con una lista de objetivos. Marcó dos mezquitas con la palabra «hoy» y otras dos con la palabra «mañana». Los grupos se llenaron sobre todo con contenido de Weerasinghe. Muchos compartieron un vídeo que había publicado en Facebook y YouTube en el que aparecía caminando por las tiendas de una población llamada Digana. Había demasiados negocios regentados por musulmanes, decía, y urgía a los cingaleses a retomar la ciudad. Los investigadores mandaron todo eso a Facebook. Nunca obtuvieron respuesta.

Vieron impotentes como cientos de cingaleses hacían publicaciones en directo desde las aldeas y ciudades cuyas calles llenaban. Los residentes colgaban banderas con imágenes de leones en las ventanas delanteras. Era un mensaje: aquí viven cingaleses. Todo el mundo sabía lo que se avecinaba. Esa noche volaron los primeros cócteles Molotov. A lo largo de tres días, las turbas dominaron las calles. Yendo casa por casa, donde vivían musulmanes, reventaban las puertas, lo saqueaban todo de arriba abajo y prendían fuego a la vivienda. Quemaron mezquitas y tiendas propiedad de musulmanes. Golpearon a personas en las calles.

En Digana, la población por la que Weerasinghe se paseaba en ese vídeo, una de las casas pertenecía a la familia Basith. Vendían zapatillas en la planta baja y vivían en el piso de arriba. La mayoría habían huido. Pero uno de los hijos mayores, Abdul, se había quedado y estaba atrapado en el piso de arriba. «Han destrozado todas las puertas de nuestra casa —dijo Abdul en un mensaje de audio que mandó a su tío por WhatsApp—. Están entrando las llamas.» Tras unos instantes, imploraba, levantando la voz: «La casa está ardiendo». Su familia no pudo llegar al domicilio. La policía no retomó el control de Digana hasta la mañana siguiente. Encontraron a Abdul muerto en el

piso de arriba.

Los dirigentes del país, desesperados por contener la violencia, bloquearon el acceso a las redes sociales. Fue una palanca que se habían resistido a accionar, reticentes a bloquear unas plataformas que algunos aún consideraban responsables de la reciente transición del país a la democracia y temerosos de que pareciera que reinstauraban los abusos autoritarios de las décadas anteriores. Dos cosas sucedieron casi de inmediato. Se detuvo la violencia; sin Facebook ni WhatsApp motivándolas, las muchedumbres enfurecidas se fueron a sus casas. Y los representantes de Facebook, tras meses ignorando ministros del Gobierno, al fin les devolvieron las llamadas. Pero no para preguntarles por la violencia, sino para saber por qué el tráfico se había reducido a cero.

Apenas unos días después, Amanda y yo llegamos a Digana, donde las cenizas aún volaban por las calles. La población, en el interior de Sri Lanka, con sus ondulantes colinas esmeralda y sus reservas naturales, estaba situada a tan solo treinta minutos de algunos de los complejos turísticos más lujosos del país. Los vecinos observaban desde puestos de té cómo un hombre llamado Fazal nos daba la bienvenida a su casa, a escasos metros del almacén de la vivienda donde su hermano, Abdul, había muerto entre las llamas. Fazal, que trabaja como imán, utilizaba Facebook para todo, lo mismo que el resto de la gente, decía. Le pregunté por la desinformación y el odio en internet, pero no pareció entenderme. Facebook sencillamente existía. Lo mismo podría haberle preguntado si la culpa del fuego era del viento. No quería presionar a un hombre que estaba de luto. Nos ofreció helado y se fue al trabajo.

Un joven vecino que se había reunido con nosotros en casa de Fazal, Jainulabdeen, nos contó, en cuanto se fue nuestro anfitrión: «Nos lo imaginábamos». Quizás para no incomodar a Fazal, había esperado a sacar el tema. Al igual que la familia Basith, Jainulabdeen también era musulmán. Pero unos vecinos cingaleses le habían avisado con días de antelación. «La mayoría lo sabían —dijo—. Lo sabían por Facebook.» Cuando le pregunté por el vídeo de Weerasinghe, el extremista de Facebook que se paseó por Digana reclamando la expulsión de los musulmanes, Jainulabdeen resopló e hizo un gesto de negación con la cabeza. «Lo conocemos —dijo—. Es de la zona.» En Facebook, Weerasinghe tenía la capacidad de moldear

la realidad para cientos de miles de personas. Pero aquí, en su ciudad natal —insistió Jainulabdeen— solo era «una persona normal». Su padre era carpintero. Las familias se conocían. Unos parientes de Jainulabdeen incluso le habían pedido a la familia de Weerasinghe que interviniera. La familia, que al parecer compartía sus preocupaciones, había prometido que hablarían con él, pero no había servido de nada. Le gustaba demasiado pasar tiempo en internet.

Una vez que se dispersaron las muchedumbres, la policía detuvo a Weerasinghe por incitación a la violencia. Facebook terminó cerrando su página. Pero el vídeo de Ampara que había inspirado tanta violencia, del musulmán inocente Farsith Atham-Lebbe presionado para que confirmase una guerra racial inexistente, siguió colgado en internet. Los investigadores continuaron enviando peticiones para que Facebook lo eliminara, y la compañía siguió negándose, ya fuera ignorando sus denuncias o respondiendo que el contenido no infringía las normas.

En cuanto a Farsith, nos enteramos de que estaba escondido en la otra punta del país. Mientras a mí me llevaban gratis a esa zona con la esperanza de reunirme con él, Amanda se dirigió de vuelta a la capital para recabar los detalles de una reunión cuya existencia había conocido gracias a una fuente. Ese mismo día, el director de políticas de Facebook en el sur de Asia, Shivnath Thukral, había llegado en avión para reunirse con ministros del Gobierno, según nos había desvelado la fuente. Ahora que Sri Lanka había apagado la máquina, por fin Facebook daba muestras de prestar atención.

Thukral era conciliador, le contó un asistente a Amanda. Admitió que Facebook no había atajado la incitación y los discursos de odio de los que le habían advertido en repetidas ocasiones. Prometió que habría una mejor colaboración. Al día siguiente, Thukral mantuvo una llamada extraoficial con representantes civiles. Reconoció que Facebook no tenía suficientes moderadores de habla cingalesa para controlar la desinformación y el odio. Una vez más, se comprometió a que la compañía contrataría más personal.

Transcurridas varias semanas, preguntamos a Facebook cuántos moderadores de habla cingalesa habían contratado. La compañía dijo solamente que habían hecho progresos. Escéptica, Amanda rastreó portales de búsqueda de empleo de países cercanos. Encontró un anuncio, en India, que ofrecía trabajo de moderador para una

plataforma que no se mencionaba, en cingalés. Con la ayuda de un traductor, llamó a la empresa subcontratada y preguntó si la oferta era para trabajar para Facebook. El responsable de contratación dijo que sí. Tenían veinticinco puestos, todos ellos vacantes desde junio de 2017, nueve meses antes. El «progreso» de Facebook había sido una mentira.

«Somos un Gobierno que llegó al poder con un mandato de libertad de expresión», le dijo a Amanda Harindra Dissanayake, un asesor presidencial de Sri Lanka. Él también utilizaba las redes sociales. Le había dolido tener que bloquear el acceso, aunque fuera solo unos días. En el mejor de los casos, dijo, las redes sociales «aportaban mayor transparencia, daban voz a personas que no tenían». Pero los meses anteriores, añadió, habían echado por tierra su fe en la tecnología que antes creía que había contribuido a traer la democracia al país. «Esa idea de que las redes sociales son una plataforma abierta y equitativa es una mentira total y absoluta —creía ahora—. No hay ningún editor; solo el algoritmo.»

Insistió en que las divisiones de Sri Lanka eran anteriores a las redes sociales. Pero alertó de que esas plataformas sacaban lo peor de una sociedad, pues amplificaban sus extremos de formas que antes nunca habían sido posibles. «No le echamos toda la culpa a Facebook —dijo Dissanayake—. Las semillas son nuestras, pero Facebook es el viento, ¿comprendes?» Su Gobierno se estaba planteando introducir regulaciones o imponer multas, dijo. Pero sabía que el poder de Sri Lanka era limitado. Solo los americanos, creía él, tenían una capacidad de influencia suficiente para forzar un cambio. «Vosotros, Estados Unidos, deberíais combatir el algoritmo. ¿Qué fuerza a Facebook a actuar, aparte de eso?»

Al día siguiente, llegué a la otra punta del país, donde un maestro local que afirmaba conocer a Farsith me guio hasta un pequeño asentamiento a algunos kilómetros de Ampara, hasta una hilera de casas de hormigón de dos habitaciones. Señaló la tercera empezando por el final.

Farsith, que me esperaba dentro, se había afeitado la barba. No para esconder su fe, dijo, sino porque incluso en esa remota aldea apenas podía salir a la calle sin que lo reconocieran. «La gente me pregunta toda clase de cosas», me contó. O le gritaban: «¡Tú eres el del vídeo!». Relató los disturbios, la confusión y el miedo que pasó, la

furia de la turba. «Pensé que ese sería el último día de mi vida», dijo. Había huido a la mañana siguiente.

Tímido, casi anñado, parecía estar en alguna otra parte. Mientras hablábamos, tenía una mano retorcida delante de su sobrina de cinco años y jugaba sin ganas. Ella le tiraba de la mano intentando que sus ojos dejaran de mirar al suelo. El padre de la niña, quien regentaba el restaurante, nos trajo plátanos y té. Los hermanos habían pedido tantos créditos para levantar el negocio, dijo, que no se habían podido permitir un seguro. Ahora no les quedaba nada aparte de las deudas.

«No sabemos qué hacer», dijo el hermano de Farsith. Quizás volverían a trabajar en la construcción en Arabia Saudí, donde habían ahorrado para el restaurante, aunque eso significaría alejarse de sus familias. «Esperamos que Dios nos guíe.»

Farsith suspiró. «No tengo ninguna intención de quedarme aquí», declaró.

Le pregunté varias veces por las redes sociales. Facebook lo había convertido en un villano nacional. Había propagado una mentira que había arruinado a su familia y quizás ahora los obligaría a separarse. Casi lo había matado. Ahora aún vivía con miedo de que apareciese otra turba furibunda incitada por la plataforma.

Pero, a pesar de todo eso, se negaba a abandonar las redes sociales. Como ahora, escondido, tenía largas jornadas sin nada que hacer, dijo, «tengo más tiempo y paso mucho más rato en Facebook».

Sus palabras me asombraron. Aunque no sintiera animadversión por la compañía cuya plataforma había trastornado la vida de su familia, le dije yo, él sabía de primera mano que no podía creer lo que veía allí.

No era que tuviese fe en que las redes sociales decían la verdad, dijo. «Pero tienes que emplear tiempo y dinero en ir al mercado a comprar un periódico. En cambio, puedo encender el móvil y leer las noticias aquí. —Levantó la mirada del suelo y se encogió de hombros—. Tanto si es cierto como falso, esto es lo que leo.»

Seguí en contacto de forma intermitente con Farsith. Su familia cayó en la pobreza. Continuó recibiendo amenazas. Alguien de Facebook se puso en contacto con él —citando el artículo que habíamos escrito Amanda y yo— para preguntarle qué había sucedido. Farsith le dijo que estaba desesperado por encontrar una forma de alimentarse. Estaba deseando trabajar. La llamada terminó y nunca

más nadie de Facebook se puso en contacto con él. Pasado un año, había ahorrado lo bastante para viajar a Kuwait, donde empezó a trabajar como jornalero. Y allí sigue.

Las nuevas campanas

AMENAZA AL ESTATUS

Durante una entrevista con Gema Santamaría, una experta en la violencia ejercida por grupos extraoficiales que había investigado extraños incidentes en su país natal, México, me di cuenta de que tardaría años en entender cómo podía estar desarrollándose ese patrón, si bien de formas menos obvias o con unos efectos menos evidentes, en todo el planeta, quizás incluso en Estados Unidos, donde solo empezaban a vislumbrarse paralelismos con el ascenso de Trump. Santamaría estaba encontrando en México los mismos tipos de brotes que habían documentado investigadores en otros países del mundo.¹Un barrio residencial de Cancún donde estalló la violencia por culpa de la desinformación en internet.²Un pueblo de familias tranquilas donde, tras crear una página de Facebook para publicar noticias de la comunidad que se convirtió en un hervidero de rumores paranoicos, ataron a un par de encuestadores itinerantes perplejos —a quienes acusaban de conspirar para extraer los órganos de niños de la zona— y les prendieron fuego.³Luego, en otro pueblo, el mismo patrón: desde los detalles del rumor hasta el método del asesinato, esta vez con la muerte de dos hombres que se encontraban en la ciudad para comprar postes para una valla.⁴

«Las redes sociales desempeñan la función que antes hacían las campanas de las iglesias —dijo Santamaría—. Esta es la forma en que la gente sabe que se va a producir un linchamiento.» Las plataformas, me contó Santamaría, reproducían ciertos viejos mecanismos mediante los cuales una comunidad terminaba incurriendo en actos de violencia colectiva. El linchamiento, cuando un grupo lleva su indignación moral hasta el extremo de herir o matar a una persona —la tiranía de los primos en acción—, es un impulso comunitario.⁵Una demostración pública de lo que les ocurre a quienes transgreden las normas de la tribu.

«Su objetivo es comunicar», dijo Santamaría en referencia a los linchamientos. Los rumores falsos que siempre se difunden antes de los actos de violencia masiva —creía la experta— eran la pista de que las redes sociales habían aprendido a reproducir ese antiguo proceso. Más que desencadenar ese sentimiento preexistente, las redes sociales lo creaban. Los rumores casi nunca eran aleatorios. «Hay una lógica detrás —dijo—. No van dirigidos a todo el mundo.» Los rumores activaban más bien un sentimiento de peligro colectivo en grupos que eran dominantes pero que sentían que su estatus estaba en riesgo: mayorías furiosas y temerosas ante cambios que amenazaban con minar su posición en la jerarquía. Como las fuerzas impersonales del cambio social son, para la mayor parte de la gente, tan inevitables como el tiempo, las redes sociales habían aparecido para proporcionar un villano más corpóreo y conquistable: las blogueras feministas, la minoría religiosa más cercana, los refugiados. «Eso es, por fin, algo que puedes controlar —dijo Santamaría—. Ante eso sí puedes hacer algo para ponerle remedio.»

En Birmania, las redes sociales dieron alas a los miedos de la mayoría budista, tradicionalmente dominante, que con la llegada de la democracia sentía que se había producido un cambio en el *statu quo* que durante mucho tiempo la había privilegiado. En India, eso le ocurría a la mayoría hindú por motivos parecidos. En 2018, periodistas de la BBC encontraron el mismo patrón en el norte de Nigeria, donde la mayoría fulani se enfrentaba a la minoría berom, todo eso en Facebook.⁶ En Estados Unidos, las redes sociales habían aprovechado las críticas de los blancos a la inmigración, el movimiento Black Lives Matter, la creciente visibilidad de los musulmanes, la recalibración cultural hacia una mayor tolerancia y diversidad. Los rumores más compartidos, señaló Santamaría, a menudo tenían que ver con la reproducción o la población. Sri Lanka y las píldoras esterilizadoras. Estados Unidos y un complot liberal para sustituir a los blancos por refugiados.

El elemento definitorio en todos esos rumores era algo más específico y peligroso que la indignación generalizada: un fenómeno denominado amenaza al estatus. Cuando los miembros de un grupo social dominante sienten que corren el riesgo de perder su posición, puede desencadenarse una reacción violenta.⁷ Empiezan a sentir nostalgia por un pasado —real o imaginario— en el que se sentían

seguros en su predominio («Make America Great Again», ‘Que Estados Unidos vuelva a ser un gran país’). Se vuelven ultrasensibles a cualquier cambio que pueda parecer vinculado con su posición: evoluciones demográficas, normas sociales cambiantes, ampliación de los derechos de las minorías. Y se obsesionan cada vez más con presentar a las minorías como algo peligroso, además de difundir historias y rumores para confirmar esa creencia. Es una suerte de mecanismo de defensa colectivo para preservar el dominio. En su mayor parte, es inconsciente, casi animal, y por tanto fácilmente manipulable, ya sea por parte de unos dirigentes oportunistas o de unos algoritmos que solo buscan obtener beneficios.

El problema no es solo que las redes sociales aprendieron a fomentar la indignación, el miedo y el conflicto tribal, sentimientos que van de la mano de la amenaza al estatus. En internet, cuando publicamos actualizaciones visibles para cientos o miles de personas, cargadas con emociones basadas en el grupo que las plataformas incentivan, «nuestras identidades de grupo son más distintivas» que nuestra identidad individual,⁸ como escribieron William Brady y Molly Crockett en su artículo sobre los efectos de las redes sociales. No es solo que nos volvamos más tribales, es que perdemos nuestro sentido del yo. Es un entorno, escribieron los dos expertos, «perfecto para el estado psicológico de la desindividuación».

La definición rápida de desindividuación es «mentalidad de turba», aunque es más común que sumarse a una muchedumbre enfurecida. Puedes desindividuarte sentándote en las gradas de un partido o cantando con el resto de los feligreses en una iglesia, lugares en los que entregas parte de tu voluntad a la del grupo. El peligro empieza cuando esas dos fuerzas se mezclan: la desindividuación, con su capacidad de eclipsar el juicio individual, y la amenaza al estatus, que puede desencadenar agresiones colectivas hasta límites terribles.

Me acordé de una conversación con Sanjana Hattotuwa, el apasionado investigador digital que había rastreado el odio virtual en Sri Lanka. «El cáncer ha crecido tanto que encuentras a personas normales y corrientes —había dicho—. Es perturbador. La radicalización se está produciendo a una edad muy temprana.» Había incluso alumnos de colegio de buenas familias —si tenían actividad en las redes sociales— que terminaban arrastrados, y su mundo y su visión del mundo quedaba definida por la amenaza al estatus que

habían visto por internet. «Esa es su iniciación a las relaciones comunitarias —dijo—. Y es odio. Es algo muy pero que muy negativo.»

Tal vez ese patrón, según el cual la amenaza al estatus se descontrola en internet, permita explicar en parte por qué, en 2016, los seguidores de Trump habían caído mucho más abajo por la madriguera digital que los demás estadounidenses. Si las redes sociales estaban hechas para activar el pánico identitario mayoritario, entonces la mayoría blanca de Estados Unidos, que estaba en retroceso —y sobre todo las personas blancas de clase obrera o sin estudios universitarios que suelen tener un mayor apego a su identidad racial y que se convirtieron en el grueso de la coalición de Trump— serían peligrosamente susceptibles de caer en el mismo patrón que había visto en Sri Lanka. Amenaza al estatus y desindividuación digital a escala nacional. En 2018, esa tribu aún no había participado, con algunas excepciones como la concentración de Charlottesville, en actos grupales directamente violentos. Pero yo me preguntaba si esa suerte de influencia de las redes sociales podía estar saliendo por otras vías, lo cual prepararía a la gente para la violencia racial de formas menos obvias pero aun así relevantes.

Pronto obtuve una respuesta. Justo cuando Sri Lanka ardió en marzo de 2018, dos científicos sociales alemanes estaban a punto de terminar un largo proyecto en el que analizaban los efectos subterráneos de las redes sociales en su país. El estudio apuntaba a una revelación estremecedora. Daba a entender que hechos como los de Birmania o Sri Lanka, lejos de ser únicos, ya se estaban produciendo en las democracias occidentales, solo que de formas más sutiles. Para entenderlo, viajé a una pequeña ciudad cerca de Dusseldorf, donde mi compañera Amanda Taub se reuniría conmigo algunos días más tarde.

INTOXICACIÓN POR IRONÍA

Durante dos días en junio de 2018, varios meses después de sacar las informaciones sobre Sri Lanka, estaba paseando por las calles adoquinadas de Altena haciendo una pregunta que generaba gestos serios y cómplices. ¿Qué le pasó a Dirk Denkhaus?

Altena, al igual que otras muchas ciudades del noroeste industrial de Alemania, estaba en declive —me explicaron habitantes de la población—, una situación que hacía que muchos jóvenes sintieran aburrimiento y desilusión. Alemania acababa de acoger a casi un millón de refugiados procedentes de zonas remotas en conflicto, lo que la mayor parte de los ciudadanos de Altena apoyaba. Pero a algunos el influjo les había parecido desorientador. Ese era el contexto, decían, para entender por qué Denkhaus, un joven bombero en prácticas al que no se había considerado ni peligroso ni politizado, había intentado incendiar una residencia para refugiados mientras dentro dormían varias familias.⁹

No obstante, las personas con las que hablé por la calle, jóvenes o mayores, citaron en repetidas ocasiones otro factor que consideraban tan importante como los demás: Facebook. Allí todo el mundo había visto rumores en las redes sociales que presentaban a los refugiados como una amenaza. Habían leído el veneno que circulaba por grupos de Facebook locales, un fuerte contraste con los espacios físicos de Altena, donde la gente saludaba con amabilidad a las familias de refugiados. En la población mucha gente sospechaba —y eso argüirían más adelante los fiscales— que Denkhaus se había aislado en un mundo virtual de paranoia racista que lo había ido transformando de forma paulatina.

Altena ejemplificaba un fenómeno del que se sospechaba desde hacía tiempo pero que en 2018 aún se había estudiado poco: que las plataformas de redes sociales hacen que comunidades enteras sean más propensas a la violencia racial. La población era uno de los más de tres mil puntos de datos de un estudio en el que se afirmaba demostrar esa idea. Karsten Müller y Carlo Schwarz, investigadores de la Universidad de Warwick, en el Reino Unido, habían recabado datos sobre todos los ataques antirrefugiados ocurridos en Alemania durante un periodo de dos años, 3.335 en total.¹⁰ Había sido un periodo volátil, pues la crisis de los refugiados en Europa había dado paso a un ascenso de la extrema derecha en política. Las meras cifras representaban una oportunidad para aislar la influencia de las redes sociales. Por cada incidente contabilizado en el estudio, los investigadores analizaron la respectiva comunidad local utilizando diversas variables clave. El nivel de riqueza. La demografía. La afiliación política. El número de refugiados. El historial de delitos de

odio.

Había un elemento que llamaba la atención. Las poblaciones con un uso de Facebook mayor que la media solían experimentar más ataques contra los refugiados. Eso se cumplía en casi todos los tipos de comunidades: grandes o pequeñas, adineradas o con dificultades, liberales o conservadoras. El repunte no guardaba correlación con un uso de internet en general; era algo específico de Facebook. Los datos se reducían a una estadística sobrecogedora: allí donde el uso de Facebook por persona aumentaba una desviación estándar por encima de la media nacional, los ataques contra los refugiados se incrementaban cerca de un 35 %. A escala nacional, estimaban los investigadores, ese efecto era responsable de hasta un 10 % de toda la violencia contra refugiados.

Los expertos a quienes pedí que revisaran los resultados los calificaron de creíbles y rigurosos. Aun así, más adelante el estudio generó críticas por florituras metodológicas. Para medir el uso de Facebook de cada población, por ejemplo, los investigadores monitorizaron una batería de indicadores, uno de los cuales era cuántos usuarios se habían unido a la página de seguidores de Nutella. Sostenían que Nutella era una marca con una popularidad universal y neutra desde un punto de vista cultural, lo cual la convertía en un indicador útil. Los que lo criticaron afirmaban que esa elección carecía de seriedad y solidez. Los investigadores corrigieron los problemas en una versión posterior. Mi interés, no obstante, no era demostrar que los cálculos del final del artículo eran correctos, sino utilizar el estudio como una hoja de ruta para analizar la influencia de Facebook. Por eso había ido a Altena, donde los investigadores habían hallado que tanto el uso de Facebook como el sentimiento antirrefugiados eran inusualmente altos y en unas proporciones que se correspondían con las proyecciones de su artículo. Tal vez Denkhuis representase una tendencia de mayor alcance.

Cuando llegaron refugiados por primera vez a esa población unos años antes, en 2015, fueron tantas las personas del sitio que se habían ofrecido voluntarias para ayudar, que Anette Wesemann —que había asumido la dirección del centro de integración de refugiados de la ciudad después de dejar la bulliciosa Hannover para vivir en un pueblo tranquilo— no daba abasto. Encontraba familias sirias o afganas acompañadas por séquitos de autoproclamados *coaches* vitales

y profesores de alemán. «Fue muy conmovedor», me dijo. Pero, cuando creó una página de Facebook para organizar eventos de voluntariado, esta se llenó de comentarios contra los refugiados de una virulencia que nunca había visto en el mundo físico. Algunas publicaciones eran amenazadoras y mencionaban a refugiados locales por su nombre. Con el tiempo, su ira resultó ser contagiosa y terminó dominando la página. Cuando mencioné las investigaciones que relacionaban Facebook con la violencia antirrefugiados, ella respondió: «Yo pondría la mano en el fuego que es así».

El sentimiento antirrefugiados es una de las expresiones más puras de amenaza al estatus, pues combina el miedo al cambio demográfico con el tribalismo racial. Aunque fueran pocos los vecinos que odiaban de verdad a los refugiados, sus publicaciones no dejaban de aumentar, premiadas por su capacidad provocativa, como el contenido antivacunas que Renée DiResta había visto que inundaba los grupos de padres. A medida que su odio conquistaba las páginas locales —lo que, como es habitual, creaba una falsa impresión de consenso—, más personas parecían unirse.

Dirk Denkhaus resultó haber experimentado un microcosmos de ese proceso. Cuando me reuní con Gerhard Pauli, el fiscal jefe de la región, que había supervisado la investigación sobre Denkhaus, sacó un archivador que contenía cientos de documentos impresos de publicaciones de Facebook y WhatsApp que la policía había extraído del móvil de Denkhaus. Su descenso hacia el extremismo, dijo Pauli, había empezado como una broma. Él y un amigo intercambiaban memes racistas, a menudo copiados de grupos de Facebook públicos, para provocarse y sorprenderse mutuamente.

«Terminaron bromeando y dirigiéndose el uno al otro como “mein Führer” y cosas por el estilo», dijo el fiscal haciendo un gesto de desconcierto con la cabeza. Con el tiempo, el sentimiento se volvió sincero. «Hay muy poca distancia —dijo Pauli— entre la broma y la realidad.» Denkhaus recorrió esa distancia en unos seis meses. «Un día le dije a su pareja: “Y ahora tenemos que hacer algo”», me contó Pauli. Esa noche él y su amigo entraron por la fuerza en el desván de un centro para refugiados y lo incendiaron presuntamente para matar a todas las personas que había dentro. Por fortuna, el incendio se apagó. La policía detuvo a los dos hombres al día siguiente.

Existe un término para el proceso que Pauli describió, cuando las

bromas virtuales se van interiorizando de un modo paulatino como algo sincero. Se llama intoxicación por ironía. Los usuarios que pasan mucho tiempo en las redes sociales a menudo dicen de sí mismos que están «intoxicados por la ironía»,¹¹ una broma sobre la insensibilización resultante de pasarte la vida absorto en subculturas de las redes sociales, donde predominan el desapego irónico, la sobreestimulación algorítmica y el humor de atreverse a ofender. En sus formas más extremas, la exposición sostenida a contenido objetable, habiendo caído en reiteradas ocasiones por madrigueras de Facebook y YouTube, puede bajar las defensas de una persona contra eso. La desensibilización hace que las ideas parezcan menos tabú o menos extremas, lo que a su vez hace que sea más fácil adoptarlas.

Ante el tribunal, el abogado de Denkhaus insistió en que, en su vida física, su cliente esa noche no había dado muestras de animadversión contra los refugiados.¹² Si bien su intención era minimizar la importancia de las redes sociales, en realidad esa observación enfatizaba su poder. En la Altena real, predominaban mayoritariamente unas normas sociales tolerantes. Pero en Facebook, un entorno cerrado con sus propias reglas morales, Denkhaus había sido arrastrado sin control hacia el extremismo.

Pauli cree que Denkhaus representaba una tendencia. El fiscal dijo que estaba «muy seguro» de que las redes sociales habían potenciado el aumento de la violencia en Altena. Unos meses después, el alcalde fue apuñalado por un hombre que afirmaba estar indignado por las políticas prorrefugiados de la población. La policía, dijo Pauli, sospechaba que existía una vinculación con las redes sociales. Las páginas de Facebook locales se habían llenado con ira contra el alcalde justo antes del ataque. No obstante, la policía no se había tomado la molestia de recabar pruebas de la influencia virtual, pues el atacante ya había confesado. Y, aunque Pauli consideraba que el gigante de Silicon Valley era una suerte de cómplice involuntario, sabía que la compañía estaba por encima de cualquier justicia que él pudiese impartir.

Su oficina dedicó más y más tiempo a rastrear la incitación en las plataformas. A él le preocupaban cada vez más, me dijo, los rumores que podían llevar a personas normales a actuar con violencia. Aunque parezca extraño, como en México e Indonesia y al parecer en todos los demás países, a menudo parecían recurrir a misteriosas amenazas

contra niños. «Tenemos un montón de situaciones en las que una persona ve a alguien fuera del jardín de infancia —dijo Pauli haciendo un gesto de incompreensión con la cabeza—. Pasados cinco minutos, se está propagando —añadió—, y en cada publicación empeora. Pasan dos horas y tienes a una turba en la calle que quiere linchar a alguien.»

LOS SUPERPUBLICADORES

Traunstein, un pueblo de montaña cercano a Austria, es en muchos sentidos lo contrario de Altena. Su economía, basada en el turismo, es próspera. En términos políticos, es de tendencia liberal. Los jóvenes son activos en la comunidad. Pero, al igual que en Altena, el uso de Facebook y la violencia contra los refugiados alcanzan unos niveles excepcionales. Llegué a Traunstein —ahora acompañado por Amanda— buscando algo concreto. Al revisar grupos locales de Facebook para encontrar a los usuarios más activos y visibles, dimos con lo que se conoce como un superpublicador, alguien que presumiblemente encarna la forma en que Facebook puede hacer que una comunidad se vuelva cada vez más extrema. Su nombre era Rolf Wassermann.

Sea cual sea la imagen que tengáis en la cabeza de un adicto a internet que vive en un sótano, Wassermann es lo opuesto. De mediana edad y con la piel bronceada, artista de profesión, con una barba entrecana y un traje negro de arriba abajo, parece salido de un anuncio televisivo de una cerveza cara. Aunque es conservador, no es radical en absoluto. Pero tiene una actividad frenética en internet, donde encaja en el perfil arquetípico del superpublicador. Publica rumores a raudales, columnas de opinión estridentes y noticias sobre crímenes cometidos por refugiados. A pesar de que ninguno de los que vi cruzaba el umbral de los discursos de odio o de las noticias falsas, en conjunto presentaban una Alemania asediada por forasteros peligrosos.

«En Facebook, es posible llegar a personas que no están muy politizadas —nos dijo mientras tomábamos un café—. Puedes crear las opiniones políticas de la gente, en Facebook.» Describió lo que, según él, era una evolución típica de las personas que conocía en la red. Empezaban no estando muy politizadas. Comenzaban a publicar con

frecuencia, quizás porque de repente tenían mucho tiempo libre, sobre los temas que salieran en su canal de noticias. Su unían a grupos de Facebook, y allí era donde a menudo los encontraba. Con el tiempo, se volvían más estridentes políticamente, dijo Wassermann. Como le había ocurrido a él.

Él prefería las redes sociales a los periódicos o la televisión, dijo, porque «Facebook es más honesto». Por ejemplo, en Facebook se había enterado, nos contó, de que el número de refugiados en Alemania y los crímenes que habían cometido eran más elevados de lo que los medios afirmaban. Y había hecho todo lo posible para amplificar esa revelación. «Las cosas que dice la gente en Facebook son, sencillamente, más ciertas», dijo. Como si se diera cuenta de la absurdidad de creerse algo así por pura fe, añadió riendo: «Vaya, supongo que lo son. No soy Dios, no lo sé».

Los usuarios hiperactivos como Wassermann tienden a ser «más sesgados, más extremos, a estar más absortos, más todo», dijo Andrew Guess, sociólogo de la Universidad de Princeton. Es un conjunto de rasgos distinto de aquellos que podríamos asociar con los adictos y usuarios tempranos de las redes sociales como Adam, el seguidor de 4chan, un perfil muy estudiado y entrevistado. Los superpublicadores son un grupo aparte, un grupo al que las plataformas han concedido una influencia excepcional. Cuando usuarios más esporádicos abren las redes sociales, a menudo lo que ven es un mundo moldeado por los superpublicadores. Las redes sociales atraen a personas con ciertos tics de personalidad que hacen que el consumo intenso sea enormemente gratificante. Su predominio, a su vez, distorsiona las normas y los sesgos de las plataformas.

Y esos rasgos y tics definitorios de los superpublicadores, cartografiados en una serie de estudios psicológicos, son negativos en términos generales. Uno es el dogmatismo, «una certidumbre relativamente inalterable, injustificada».¹³ Las personas dogmáticas tienden a ser estrechas de miras, agresivas y ruidosas. Otro rasgo: un narcisismo ostentoso, definido por sentimientos de superioridad y privilegio innatos.¹⁴ A los narcisistas los consumen los deseos de admiración y pertenencia, lo cual convierte la repuesta instantánea y el gran público de las redes sociales en algo casi irresistible. Esa necesidad se ve agudizada por la autoestima de los superpublicadores, inusualmente baja, agravada por las propias plataformas.¹⁵ En un

estudio se concluyó sin más que «la hostilidad política virtual la llevan a cabo individuos que están predispuestos a ser hostiles en todos los contextos».¹⁶Y eso lo confirmaron experimentos neurológicos:¹⁷los superpublicadores se sienten atraídos y recompensados por el poder social negativo, un término clínico para referirse al placer obtenido al infligir de forma deliberada sufrimiento emocional a otros seres. Además, al utilizar más las redes sociales y al ser recompensados por ello con una mayor visibilidad, los superpublicadores arrastran las plataformas hacia esas tendencias definitorias del dogmatismo, el narcisismo, la exaltación y la crueldad.

En un test involuntario sobre eso realizado en 2015, Ellen Pao, que aún estaba al mando de Reddit, intentó algo sin precedentes: en lugar de promover a los superusuarios, Reddit prohibiría a los más tóxicos. De las decenas de millones de usuarios —concluyó su equipo—, solo unos 15.000, todos hiperactivos, generaban la mayor parte del contenido de odio. Expulsarlos, pensaba Pao, podía cambiar Reddit de arriba abajo. Tenía razón, como halló un análisis externo.¹⁸Con la eliminación de ese minúsculo porcentaje de usuarios, los discursos de odio disminuyeron un asombroso 80 % entre quienes permanecieron en la plataforma. De la noche a la mañana había cambiado el comportamiento de millones de personas. Era un éxito infrecuente en el abordaje de un problema que en otras plataformas más grandes, que no siguieron los pasos de Reddit, no hizo más que agravarse. No tenían interés alguno en eliminar a sus usuarios más activos, y menos aún en reconocer que podía darse el caso de que alguien pasase demasiado tiempo conectado.

¿Podían los superpublicadores alterar no solo lo que aparecía en los canales de noticias de los usuarios, sino su percepción de lo que está bien y lo que está mal? Le hice esta pregunta a Betsy Levy Paluck, que había obtenido una «beca para genios» de la Fundación MacArthur por sus trabajos sobre cómo las normas sociales influyen en el comportamiento.¹⁹Yo esperaba que citase sus investigaciones sobre, por ejemplo, la violencia comunitaria en Ruanda. Pero ella quería hablar del acoso escolar. Los niños acosan o no —descubrió Paluck en una larga investigación—²⁰basándose, en buena medida, no en si prevén que los castigarán o si piensan que la víctima se lo merece, sino en si lo ven como algo moralmente aceptable. El acoso escolar o era permisible, o estaba incluso justificado, o era algo

erróneo; y ese barómetro interno era lo que más importaba. Pero ¿cómo se calibra nuestro barómetro moral? Nos gusta pensar que seguimos un código moral innato, derivado de unos nobles principios, la experiencia vivida y los consejos de una persona mayor de confianza. Pero, en realidad, se ha descubierto en varios estudios que nuestro sentido de lo correcto o lo incorrecto está muy influenciado — aunque sea de forma inconsciente— por lo que creemos que piensan nuestros iguales: se trata de una moralidad por consenso tribal, guiada no por un ángel bueno o un poder superior, sino por la sumisión por supervivencia a la tiranía de los primos.

En un experimento que se hizo en una zona rural de México,²¹ los investigadores crearon una radionovela cuyo argumento convencía de que estaba mal la violencia doméstica contra las mujeres. En algunas áreas, hicieron que la gente escuchase la radionovela en privado en sus casas. En otras, la emitieron por unos altavoces del pueblo o en encuentros comunitarios. Los hombres que la escucharon en casa mantuvieron la propensión a ejercer violencia doméstica que tenían antes. Pero entre los que la escucharon en ambientes grupales disminuyó de manera significativa la propensión a perpetrar agresiones. Y no por una presión percibida. Sus creencias internas habían cambiado, se habían vuelto contrarios moralmente a la violencia doméstica y favorables a la igualdad de género. La diferencia estaba en ver a sus iguales consumir la radionovela. El impulso de conformidad —el mismo que había llevado a los primeros usuarios de Facebook a enfurecerse por el canal de noticias— puede calar hasta alcanzar la médula moral de nuestro yo más interior.

La mayoría de las veces, deducir las opiniones morales de nuestros iguales no es tan sencillo. Por eso usamos un atajo. Prestamos especial atención a algunas personas a quienes consideramos influyentes, tomamos pistas a partir de su comportamiento y suponemos que eso refleja las normas del grupo en conjunto. Las personas a las que elegimos como indicadores morales se conocen como «referentes sociales». Así pues, la moralidad es «una suerte de labor de percepción», me contó Paluck. «¿Qué individuo de nuestro grupo sobresale para nosotros? ¿A quién repescamos de nuestros recuerdos cuando pensamos en lo que es habitual o deseable?»

Para poner eso a prueba, Paluck mandó a su equipo a cincuenta y seis colegios para que identificasen qué estudiantes eran influyentes

entre sus compañeros, además de qué alumnos consideraban que el acoso escolar era aceptable moralmente.²² Luego eligió a entre veinte y treinta alumnos de cada centro que parecían cumplir ambos requisitos: esos eran, presumiblemente, los estudiantes que desempeñaban el rol más importante en inculcar normas sociales favorables al acoso escolar en sus comunidades. Se les pidió que condenaran en público el acoso escolar. No se les obligó, solo se les pidió. El suave empujoncito a ese minúsculo grupo fue transformador. En los indicadores psicológicos se vio que miles de estudiantes desarrollaron una oposición interna al acoso escolar y sus brújulas morales viraron hacia la compasión. Los informes disciplinarios por acoso escolar disminuyeron un 30 %.

Las redes sociales nos colocan a todos en una versión del experimento de Paluck en colegios. Pero en internet nuestros referentes sociales —las personas puestas de forma artificial en nuestro campo de visión moral— son los superpublicadores. No porque sean convincentes, ambles o importantes, sino porque hacen subir la participación. Eso era algo exclusivo de plataformas como Facebook, dijo Paluck. Cualquiera que pasase mucho tiempo en el canal de noticias podía ser influyente. «En la vida real, algunas personas hablan mucho pero no son las más escuchadas. Pero Facebook —dijo la investigadora— te las pone delante todo el tiempo.»

Y las redes sociales no solo ponen superpublicadores a tu alrededor. Muestran sus mensajes en enormes foros públicos, donde sabes que todo el mundo también los ve, como los altavoces de los pueblos mexicanos que demostraron un gran poder para alterar una comunidad en poco tiempo. En Alemania, las redes sociales parecían haber propulsado una clase de superpublicadores como Wassermann que daba a usuarios de toda la plataforma la impresión de que las normas sociales eran más hostiles contra los refugiados y más conspirativas de lo que eran en realidad. Era el embrollo que hubo en Facebook en 2006 «Contra el canal de noticias», ahora elevado a la psique política de todo un país y dirigido contra millones de los residentes más vulnerables. Aunque ninguno de esos superpublicadores apoyaba de forma explícita la violencia, dijo Paluck, el efecto acumulado de sus mensajes antirrefugiados y contrarios al Gobierno hacía probable que la violencia extraoficial

fuese tolerada o incluso fomentada.

Esa tarde, en un acto social en Traunstein, a una profesora llamada Natascha Wolff le llamó la atención que yo preguntase por las redes sociales. Wolff daba clase en un centro de formación profesional, explicó, con una mezcla de alumnos nacidos en Alemania y nacidos en el extranjero. Durante los meses anteriores, los niños alemanes habían virado, casi de manera uniforme, hacia una estridente hostilidad antirrefugiados que ella nunca había visto antes. La docente sabía que eso probablemente se debía a muchas razones. Sin embargo, cuando preguntaba dónde habían encontrado las estadísticas inventadas o las afirmaciones llenas de odio que se repetían alarmados unos a otros, siempre obtenía la misma respuesta: en Facebook.

Cualquier rumor o chismorreio de Facebook menospreciando a los extranjeros, dijo la profesora, «se extiende con rapidez. La gente ve confirmado su punto de vista». Y añadió, moviendo los brazos arriba y abajo imitando a alguien aporreando un teclado: «No es más que: “me gusta”, “me gusta”, “me gusta”». Si ella ponía en duda alguna afirmación falsa, siempre le contestaban lo mismo: «Todo el mundo sabe que es verdad». Pero a menudo los estudiantes también se equivocaban en eso; en Traunstein muchas personas no se creían los rumores falsos. A Wolff le preocupaba que esa burbuja de Facebook, el falso consenso comunitario, tuviese consecuencias. A sus estudiantes refugiados les echaban café por encima en la calle o les tiraban basura desde un coche. Unos actos violentos despreocupados y a plena luz del día que uno solo hace sabiendo que se los tolerarán.

La violencia originada en las redes sociales se había vuelto tan habitual que la policía había empezado a tratar las plataformas como una amenaza constante para la seguridad pública. «Facebook no es solo como un tablón de anuncios donde la gente cuelga cosas y los demás las leen —nos contó al día siguiente el inspector de la policía municipal Andreas Guske mientras tomábamos un café—. Facebook, con su algoritmo, influye en los usuarios.» Guske, un detective veterano, con el pelo entrecano, empezó a tomarse en serio las redes sociales como una amenaza en 2015, durante una cumbre del G7 organizada cerca de allí. Cuando los manifestantes entraron en tromba, se dio cuenta de que las plataformas se llenaban de rumores, algunos de los cuales arrastraban a las multitudes hacia histerias paranoicas. Al año siguiente, pareció que los ataques contra los

refugiados aumentaban en correlación con los discursos de odio por internet. Guske reestructuró el equipo que supervisaba las comunicaciones del departamento para contrarrestarlo, tanto en el mundo virtual como en el físico. Se veían a sí mismos como unos trabajadores sanitarios, vacunando comunidades contra la desinformación viral y sus consecuencias.

Guske me habló de un caso reciente en el que Facebook había bullido con la información de que, en un pueblo cercano a Traunstein, un grupo de refugiados musulmanes había llevado a una niña de once años hasta un paso subterráneo para peatones y la había violado. El rumor, aunque era falso, provocó oleadas de indignación, pues Facebook lo difundió por toda Alemania. Cuando la policía desmintió la historia, los usuarios insistieron en que los políticos les habían dado la orden de encubrir el caso. El equipo de Guske descubrió que los rumores se habían iniciado cuando la policía detuvo a un inmigrante afgano acusado de manosear a una chica de diecisiete años. A medida que los usuarios habían ido pasándose la información, algunos habían añadido detalles que escandalizaban o indignaban, lo cual había alejado las versiones de la verdad. De un agresor se pasó a varios. Un manoseo se convirtió en una violación. Y la víctima adolescente al final era una niña.

La policía publicó comunicados en Facebook y Twitter en los que desmentía el rumor reconstruyendo cómo se había propagado. Guske creía que, si la policía podía mostrar cómo las plataformas distorsionaban la realidad, convencería a los ciudadanos de que rechazaran lo que hubieran visto en ellas. Pero también sabía que en las redes sociales una verificación de los hechos serena nunca tendría tanta visibilidad como un rumor obsceno. Así pues, su equipo identificó a personas de la zona que habían compartido el rumor durante los primeros días de su difusión y se presentaron en sus domicilios para demostrarles con pruebas que habían divulgado una información falsa. Los alentaron a retractarse públicamente de sus afirmaciones, con la esperanza de convertir los propios sistemas de promoción de las plataformas en armas contra la desinformación. Todos menos uno eliminaron o corrigieron sus publicaciones como Guske les había pedido. Pero nunca pudieron seguir el ritmo de las plataformas, cuya capacidad intoxicadora, temía Guske, no hacía sino acelerarse. Y lamentaba que Facebook, en ese momento una compañía

con un valor de 500.000 millones de dólares, cediese a unos departamentos de policía saturados la tarea de gestionar los riesgos que ellos creaban. «Cuesta impedir las noticias falsas, porque una vez que Facebook las propulsa... —dijo dejando la frase a medias y negando con la cabeza—. ¿Qué más puedes hacer?»

Esa tarde, mientras Amanda entrevistaba a personas de la ciudad sobre las redes sociales, yo me reuní en un parque cercano con una joven que había asistido al centro de formación profesional de Wolff. Vino con una amiga, que trajo a su hijo. Ambas mujeres, educadas pero cautas, se describieron como personas no muy politizadas. Ninguna de las dos leía las noticias salvo por lo que veían en Facebook, plataforma a la que entraban con frecuencia. Les pregunté qué opinión tenían de los refugiados, y a partir de entonces solo quisieron hablar de eso. Que los refugiados eran agresivos, que violaban y que muchos simpatizaban con extremistas, dijeron las jóvenes. Me contaron historias escabrosas e inverosímiles sobre crímenes cometidos por refugiados que el Gobierno encubría. Todo aquello lo habían leído en Facebook, el sitio donde a menudo debatían sobre «la cuestión de los refugiados», dijo una de ellas.

Traunstein es de tendencia liberal, pero está dividido políticamente, y le pregunté a la mujer si alguna vez se enzarzaba en disputas sobre los refugiados en internet. La pregunta pareció confundirla. «Todo el mundo tiene esta opinión», dijo. Su burbuja, donde el miedo era unánime, se había vuelto su realidad. Ella, al igual que Wassermann y sus amigos virtuales, así como los demás estudiantes de Wolff, además de las personas a las que Guske había suplicado que retirasen sus falsedades racistas, eran la masa sumergida de un iceberg de radicalización en las redes sociales que afectaba a la sociedad entera. Denkhaus, el bombero incendiario, era solo la punta del iceberg. Había otros muchos alemanes que también se habían vuelto más xenófobos, más conspiratorios y más nacionalistas. La mayoría nunca emplearían la violencia. Pero su deriva colectiva tenía unas consecuencias más profundas, que estaban alterando de forma invisible las costumbres y la política de la sociedad. En un país rico y democrático como Alemania, el resultado quizás no fuera tan obvio como un linchamiento o unos disturbios. Podía ser peor. El centro político del país se estaba desmoronando. La extrema derecha alemana iba en aumento.

«A uno de los alumnos de mi colegio lo mandaron de vuelta a África», dijo la mujer con gesto de aprobación. La deportación se había producido por un error en sus documentos de inmigración. «Deberían mandarlos de vuelta a todos.»

APAGONES

Los investigadores alemanes de la Universidad de Warwick sabían que había un elemento de su teoría —la causalidad— que requería una especial atención. ¿Podía demostrarse que el uso de Facebook y la violencia antirrefugiados aumentaban en paralelo justo porque lo primero causaba lo segundo? Se les ocurrió la idea de analizar todos los apagones de internet significativos que se habían producido durante el periodo que cubría el estudio. La infraestructura de internet alemana tiende a ser local, lo que hace que los apagones sean habituales pero aislados. Cada uno de ellos era una oportunidad para poner a prueba la causalidad: si privar de Facebook a una comunidad de repente hacía descender la violencia de los vecinos de la zona contra los refugiados, eso indicaría que Facebook era el motor de algunos ataques.

En efecto, cuando el acceso a internet fallaba en un área con un alto uso de Facebook, los ataques contra los refugiados disminuían de forma significativa. Sin embargo, no tenía lugar la misma caída cuando áreas con un elevado uso de internet pero solo con un uso medio de Facebook sufrían un apagón, lo cual daba a entender que el efecto suscitador de violencia era específico de las redes sociales y no del uso de internet en general. Y la violencia disminuía en el mismo porcentaje —un 35 %— en el que el estudio había sugerido que Facebook fomentaba esos ataques. Los investigadores insistían en que eso no era definitivo, solo un ejercicio para comprobar sus conclusiones. Pero era una indicación llamativa de que estaban desentrañando algo. Y era una oportunidad para plantearse, con un rigor que no podían ofrecer bloqueos excepcionales como el de Sri Lanka, qué sucede cuando desaparecen las redes sociales.

«El mundo se ha empequeñecido, han cambiado muchas cosas», dijo Stefania Simonutti al recordar el apagón que había afectado el barrio residencial de Berlín en el que vive durante entre varios días y

algunas semanas, dependiendo del bloque. El barrio, Schmargendorf, parece un lugar a salvo de las fuerzas del odio. Familias diversas de clase media pasean por avenidas con *boutiques* y mercados de agricultores con productos de calidad. Pero ahí el uso de Facebook es elevado. Y también los ataques contra refugiados. Excepto durante el apagón.

Simonutti, al preguntarle cómo lo había sobrellevado, abrió la boca y se puso las manos en las mejillas simulando un grito. Había perdido el contacto con sus familiares que viven en el extranjero y con las noticias, para las que confiaba solo en las redes sociales. «Muchas personas mienten y se inventan cosas en los periódicos —dijo—. Pero con internet, yo puedo decidir por mi cuenta qué creer y qué no.» Forzada a renunciar a las conspiraciones de las redes sociales que le gustaba seguir por internet, dijo, llenó el tiempo libre relajándose con la familia.

Todo el mundo parecía recordar el apagón. A Esperanza Muñoz, una mujer alegre y pecosa que se había trasladado allí desde su Colombia natal en los años ochenta, le había parecido relajante. Socializó más con los vecinos y siguió menos las noticias. Su hija, estudiante de Medicina, dijo que no se había dado cuenta de cuánta ansiedad le causaban las plataformas hasta que no estuvo algunos días sin ellas. El apagón, dijo, le había hecho entender hasta qué punto «cuando las noticias se propagan por Facebook, están hechas de un modo más provocador». Su madre estaba de acuerdo. Varias semanas antes, durante la celebración de unas elecciones en Colombia, dijo, su canal de noticias, dominado por compatriotas colombianos, se había llenado de disputas e indignación partidista; y, como si siguiera un guion, con comentarios que sembraban el miedo sobre los refugiados.

En abril de ese año, Zuckerberg había concedido una entrevista al jefe de redacción de *Vox*, Ezra Klein, quien lo presionó por el genocidio de Birmania. Como prueba de los progresos de Facebook, Zuckerberg dijo que, en el punto álgido del baño de sangre, el equipo de seguridad de la compañía había identificado usuarios de Birmania que incitaban a la violencia por Facebook Messenger. «Ahora, ante un caso así, nuestros sistemas detectan que eso está ocurriendo. Impedimos que esos mensajes circulen —afirmó—. Pero eso es, sin duda, algo a lo que le prestamos una gran atención.»²³

Tras la publicación de la entrevista, organizaciones pro derechos

humanos birmanas le contestaron con una iracunda carta abierta. En realidad, decían, eran ellos —y no Facebook— quienes habían encontrado los mensajes en cadena que fomentaban la violencia. Y, como no disponían de las herramientas internas de Facebook para monitorizar las plataformas de forma automatizada, habían logrado detectarlos solo gracias a lo que, según ellos, era el engorroso y tristemente insuficiente método de la búsqueda y captura manual. Aun así, las organizaciones se habían visto obligadas a bombardear Facebook con advertencias durante días antes de que por fin alguien de la compañía actuase. Pero ya era demasiado tarde. Los usuarios, al parecer actuando a raíz de esos mensajes virales, ya se habían organizado y habían ejecutado tres ataques separados, uno de los cuales consistió en intentar incendiar una escuela. El episodio, decían las organizaciones, subrayaba la «confianza excesiva» de Facebook «en terceros, lo que demuestra la falta de un auténtico mecanismo para contener el aumento de las situaciones de emergencia, reticencia a colaborar con los agentes locales en torno a soluciones sistémicas y falta de transparencia».²⁴ Zuckerberg envió a las organizaciones un correo disculpándose,²⁵ aunque solo por no haberlas citado por su nombre, lo cual —según destacaron los miembros de las organizaciones en su respuesta— no era su preocupación principal.

Ese agosto, las Naciones Unidas publicaron su informe oficial sobre el genocidio.²⁶ Calificaban el papel de las redes sociales, en concreto de Facebook, de «significativo». Pero Facebook aún se negaba a compartir sus datos con los inspectores de la ONU, decían ellos mismos, lo que suponía un obstáculo para desentrañar cómo se había producido el genocidio y, por tanto, cómo impedir otro. «No puedes chasquear los dedos y resolver esos problemas —dijo Zuckerberg un mes después—. Requiere tiempo contratar personal y formarlo, y crear los sistemas que puedan etiquetar contenido para que ellos lo revisen.»²⁷ Pero desde luego, tanto en Birmania como en Sri Lanka, Facebook no había reaccionado a los avisos de violencia inminente con la introducción de nuevas salvaguardias ni la contratación de moderadores, sino con meses de inacción. Ahora, una vez más, nada parecía cambiar, me contó una organización dedicada a la monitorización digital con sede en Birmania. Facebook había pedido al grupo que examinase el aumento de la incitación virtual y otros peligros. Pero la compañía ignoró en su mayor parte los

informes de la organización, sin importar lo urgentes que fueran. Facebook, creían, los había contratado como un gesto de cara a la galería.

Adam Mosseri, el ejecutivo que había supervisado el poderosísimo canal de noticias durante las matanzas de Birmania y Sri Lanka, fue ascendido a vicepresidente de Instagram y luego a presidente. A Jen Weedon, la gerente de políticas de seguridad de Facebook que no había respondido a las advertencias de los investigadores de Sri Lanka acerca del inminente derramamiento de sangre, también la ascendieron. Ese año, los ingresos de la compañía alcanzaron la cifra récord de más de 55.000 millones de dólares, casi un 40 % más que el año anterior.²⁸

«El problema viene del modelo de negocio», me contó ese año Hany Farid, un informático de la Universidad de California en Berkeley que había sido asesor de Gobiernos y organizaciones de derechos humanos sobre los peligros incipientes de la web social. «Cuatrocientas horas de YouTube subidas cada minuto. Mil millones de subidas a Facebook al día. Trescientos millones de tuits al día. Esto viene a ser un caos —dijo—. Las empresas tecnológicas, yo ni siquiera diría que se durmieron a la hora de hacer su trabajo. Yo diría que tenían los ojos bien abiertos. Creo que sabían exactamente lo que estaban haciendo. Sabían el veneno que había en la red. Sabían que tenían un problema. Pero lo único que les importaba era el crecimiento agresivo. Ahí empezó el problema.»

Farid tomó aire y volvió al asunto por el que le había llamado, una tecnología especializada que usaban las plataformas. Pero al cabo de un rato, cuando estaba terminando una explicación técnica, volvió a subir la voz al dar con una referencia a YouTube. «YouTube es la peor», afirmó. De las que consideraba que eran las cuatro compañías tecnológicas principales —Google/YouTube, Facebook, Twitter y Microsoft—, la mejor al gestionar lo que él llamaba «el veneno» era, según él, Microsoft. «Y tiene sentido, ¿a que sí? No es una empresa de redes sociales —dijo—. Pero YouTube es la peor en estos asuntos», repitió.

Había sido un año de escándalos y polémicas en torno a Facebook, considerada por lo general como la plataforma más influyente. Pero la advertencia de Farid tenía sentido porque, ya mientras investigaba los efectos de Facebook en Sri Lanka y Alemania,

había oído eso mismo por parte de expertos digitales, organizaciones de derechos humanos y otras personas: presta atención a YouTube. «YouTube es la que más se pasó por alto en 2016 —tuiteó un año después de las elecciones Zeynep Tufekci, socióloga de la Universidad de Carolina del Norte—. Sus algoritmos de búsqueda y recomendación son motores de desinformación.» Más adelante lo calificó como «uno de los instrumentos de radicalización más potentes del siglo XXI». Danah Boyd, fundadora de un laboratorio de ideas sobre tecnología, compartía esa idea, y le dijo a mi compañera Amanda: «YouTube es tal vez la plataforma más preocupante que existe en la actualidad».

Cada vez más aparecían historias sobre casos extraños y desestabilizantes —el ascenso de un grupo de odio, otro peligroso rumor médico, un chaval solitario que perpetraba una matanza— que mencionaban YouTube. Apenas había terminado de transcribir mis notas de Alemania cuando, unas semanas después de mi conversación con Farid, sucedió algo en la plataforma que dejó claro al instante por qué el experto había lanzado su advertencia.

La madriguera

EL DISTURBIO DE YOUTUBE

Grupos neonazis llevaban dos días ocupando las calles de su ciudad de forma intermitente cuando Sören Uhle, un político municipal esbelto y con gafas, empezó a recibir extrañas llamadas de periodistas. Por lo que él sabía, dos refugiados de Oriente Medio habían apuñalado a un hombre de la población durante una pelea y lo habían matado, lo cual los grupos de extrema derecha habían aprovechado para animar a que la gente acudiera en masa a su ciudad. Ahora, le contaban los periodistas, resultaba que en realidad los refugiados habían matado no a uno sino a dos hombres. También habían acosado sexualmente a una mujer de la ciudad; las víctimas de los refugiados habían muerto al intentar protegerla. ¿Podía Uhle hacer alguna declaración sobre el suceso? ¿Y podía explicar por qué los políticos estaban pagando en secreto a habitantes de la ciudad para que asistiesen a una contramanifestación que se celebraría dentro de poco?

Uhle estaba desconcertado. Esas revelaciones eran todas falsas. «Aquello era algo nuevo —me dijo—. Nunca me había pasado que medios convencionales, grandes periódicos alemanes y cadenas de televisión me preguntasen por noticias y propaganda falsas que se habían vuelto claramente tan omnipresentes que la gente se las tragaba.»¹ Eso era agosto de 2018. Las muchedumbres que se estaban congregando en Chemnitz, su ciudad de 250.000 habitantes situada en el este de Alemania, se habían organizado en las redes sociales, eso lo sabía. Quizás la desinformación también.

En Berlín, subiendo por la autopista hacia el norte, un investigador digital llamado Ray Serrato estaba llegando a la misma conclusión. Como todo el mundo en Alemania, había seguido con asiduidad las noticias sobre los disturbios: una demostración de fuerza neonazi surgida de la nada tan impresionante que la canciller Angela Merkel la había condenado. Luego el tío de su esposa le enseñó un

extraño vídeo de YouTube. Dos hombres de mediana edad, uno con rastas y un gorro negro, decían a cámara que los alborotadores no eran neonazis sino refugiados musulmanes. El vídeo, publicado por un oscuro grupo radical, era inconexo y de producción barata. Aun así, tenía casi medio millón de visualizaciones, mucho más que cualquier vídeo informativo sobre los disturbios. ¿Cómo era eso posible?

Movido por la curiosidad, Serrato aplicó una serie de técnicas que utilizaba en sus trabajos de monitorización de los discursos de odio en internet en Birmania para una organización de control de la democracia.² Empezó con una decena de vídeos recientes sobre Chemnitz; luego, sobre cada uno, analizó las recomendaciones de YouTube sobre qué ver a continuación. Luego hizo lo mismo con respecto a esos vídeos, y así sucesivamente. Su labor puso al descubierto una red de unos 650 vídeos: el ecosistema cultivado por YouTube de contenido sobre Chemnitz. De forma perturbadora, las recomendaciones de YouTube se agrupaban alrededor de un puñado de vídeos conspirativos o de extrema derecha. Eso hacía pensar que cualquier usuario que entraba en la red de vídeos de Chemnitz —pongamos por caso, buscando informaciones actualizadas o viendo un vídeo que le hubiese mandado un amigo— sería arrastrado por el algoritmo de YouTube hacia contenido extremista. Al preguntarle cuántos pasos harían falta, de media, para que un usuario de YouTube que encontrase un vídeo con información sobre Chemnitz terminase viendo propaganda de extrema derecha, Serrato contestó: «Solo dos. —Y añadió—: En el segundo, ya estás metido de lleno en la derecha alternativa».

Las recomendaciones casi nunca llevaban a los usuarios de vuelta a noticias de medios convencionales o a contenido liberal o apolítico de cualquier tipo. Una vez que uno estaba entre extremistas, el algoritmo tendía a mantenerlo allí, como si esa hubiera sido la destinación desde el principio. Incluso llevaba desde vídeos de Chemnitz hasta otras cuestiones de la extrema derecha no relacionadas —nacionalismo blanco, conspiraciones antisemitas—, del mismo modo que Facebook había conducido a Renée DiResta de páginas antivacunas a causas radicales totalmente desvinculadas. Un vídeo típico afirmaba que Trump era un peón de la familia de banqueros Rothschild. Aunque Serrato consideraba los vídeos aberrantes y peligrosos, admitió que tenían algo que era difícil de

desactivar. «Ese es el objetivo de YouTube —dijo—. Yo sigo viendo cosas en la plataforma, los anuncios cumplen su función. Y funciona.»

Ese efecto —me di cuenta trabajando con Katrin Bennhold, jefa de la oficina en Berlín del *New York Times*— había contribuido a generar el caos de Chemnitz. Poco después del apuñalamiento, algunos oscuros *youtubers* de extrema derecha habían subido vídeos sobre el incidente. Uno, un bloguero llamado Oliver Flesch, solo contaba con 20.000 suscriptores. Tenía escaso seguimiento y hacía poca promoción más allá de su burbuja ideológica. Sin embargo, sus vídeos sobre Chemnitz acumularon cientos de miles de visualizaciones, gracias a la fuerte promoción que les dio el motor de recomendaciones de YouTube.

Serrato descubrió que a los usuarios que veían algo sobre Chemnitz en YouTube —por ejemplo, el vídeo de una noticia— enseguida les aparecía recomendado el canal de Flesch. Este subió 14 vídeos sobre ese tema, todos los cuales salían en las recomendaciones de YouTube, lo que inundaba la plataforma de las mismas falsedades racistas sobre las que le preguntarían a Sören Uhle más adelante. Otros canales de extrema derecha o conspiraciones rápidamente agarraron la versión de los hechos de Flesch, lo cual convirtió una pelea callejera aislada en un relato sobre que la virtud de los blancos estaba en peligro. El algoritmo de YouTube también propulsó esos canales.

Incluso a los alemanes que buscaban noticias de Chemnitz en Google se les redirigía a conspiracionistas de YouTube. Google a menudo ofrece vídeos de YouTube cerca de los primeros resultados de búsqueda, un acto de sinergia corporativa diseñado para incrementar los ingresos.³ Eso significa que las prácticas de YouTube no se limitan a YouTube; puesto que Google domina las búsquedas de internet, estas prácticas influyen en cómo casi todo el mundo que navega por la red encuentra noticias e información y accede a ese contenido.

A medida que YouTube y Google encaminaban más alemanes a vídeos sobre Chemnitz repletos de falsedades, creció el interés por la ciudad, también entre muchas personas ajenas a la extrema derecha. Las voces de YouTube que recibían toda esa atención animaron a sus seguidores, en rápido aumento, a demostrar su apoyo a la víctima del apuñalamiento yendo a Chemnitz. Dicen habitantes de la ciudad que, en los días previos al brote de violencia, las teorías de la conspiración

se hicieron extrañamente habituales, y se susurraban en bares y dispensadores de agua. Luego llegaron las multitudes, deseosas de recuperar la ciudad tomada por los extranjeros. Pronto empezaron los disturbios, con el saqueo de tiendas y enfrentamientos con la policía. Muchos de los alborotadores atribuían a YouTube el hecho de estar allí.

Era el mismo estallido de Sri Lanka, pero ahora en el corazón de Europa. Sin embargo, había una diferencia importante. En Sri Lanka, las redes sociales habían radicalizado a un grupo social del mundo real con una fuerte identidad, los cingaleses. En Alemania, por el contrario, los alborotadores de Chemnitz eran algo nuevo. Había, sin duda, neonazis curtidos entre la multitud, pero muchos no pertenecían a ningún grupo o causa concreta. YouTube, en lugar de activar una comunidad preexistente con una identidad preexistente, había creado una de la nada. Había construido la red con sus sistemas, la había entrelazado con una realidad y unas creencias compartidas y luego la había entregado al mundo, todo ello en cuestión de días. Eso era algo mucho más profundo incluso que la violencia extraoficial inspirada por Facebook que yo había visto en el país tan solo unos meses antes.

Si había ocurrido con tal facilidad en Chemnitz, ¿estaba ocurriendo también en otras partes? En efecto, un investigador que por casualidad era originario de Alemania estaba demostrando en ese momento que ese mismo proceso llevaba meses produciéndose en Estados Unidos. Eso había acelerado el ascenso de una aterrorizadora nueva extrema derecha y la había mezclado con las tendencias más extremas y peligrosas de las redes sociales. El hombre había empezado a investigar ese fenómeno en 2015, lo había seguido durante algunos de los choques políticos más notables de los años posteriores hasta que, a lo largo de varios meses de 2018, presentó pruebas de lo que antes solo se había sospechado: que YouTube estaba haciendo con la política estadounidense lo que había hecho con Chemnitz, es decir, inundarla con una derecha desbocada de su propia creación algorítmica. Pero Chemnitz y otros sucesos perturbadores de ese año resultarían ser solo un eslabón más en la cadena de descubrimientos de ese experto, los cuales presagiaban acontecimientos mucho más grotescos y mortíferos.

Jonas Kaiser echó un vistazo a las profundidades de su primera madriguera de YouTube durante una pausa entre las sesiones de investigación que hacía para un doctorado que, de entrada, poco tenía que ver con las redes sociales en Estados Unidos. Estaba estudiando el escepticismo sobre el cambio climático en su Alemania natal, lo cual se consideraba una especie de misterio. A diferencia de Estados Unidos, en Alemania no había ningún gran partido ni figura política que fueran escépticos con el cambio climático, por lo cual Kaiser estaba perplejo ante la existencia de personas que dudaban. «Parecía, más que nada, una comunidad muy rara —me dijo—. Me preguntaba: pero ¿cómo?, ¿por qué?»

La vida de los estudiantes de posgrado puede llegar a ser agotadora. A Kaiser, larguirucho y calvo, que parece una escobilla con gafas y jerséis de cuello de pico, le gustaba relajarse con vídeos de YouTube de competiciones de videojuegos. Como de niño vivía en un pueblo remoto, estaba conectado con amigos mediante los videojuegos por internet, pero el portátil que tenía cuando estudiaba el posgrado no estaba equipado para ello, así que, en lugar de eso, veía vídeos. Un día, recuerda Kaiser, «el Gamergate estaba por todas partes» y saturaba los canales de *gaming* de YouTube. Hasta entonces, él nunca había oído hablar de *Breitbart* o de Milo Yiannopoulos, dijo. La plataforma empezó a arrastrarlo hacia otras comunidades que también habían adoptado la causa del Gamergate: *youtubers* favorables al ateísmo, *youtubers* de ciencia y *youtubers* de la derecha alternativa. «Percibí claramente un cambio: de repente nuevas comunidades formaban su identidad en torno a aquello —dijo—, en torno a la misoginia y la desinformación.» Guiando a los usuarios hacia esas facciones y reenviándolos de una a otra, el algoritmo parecía estar uniéndolas todas.

Fue un momento de epifanía para Kaiser, que vio un paralelismo con los negacionistas del cambio climático alemanes. «Todo es muy pequeño y troceado —dijo en referencia al movimiento de los escépticos del cambio climático de su país—. En realidad, el único lugar donde podían intercambiar sus pensamientos y fusionarse y encontrar aliados era internet.» Esos grupos no reflejaban comunidades del mundo real de un tamaño significativo, se percató el

investigador.⁴Eran nativas de la red; y, por consiguiente, estaban moldeadas por los espacios digitales que los habían nutrido. Los escépticos climáticos se reunían sobre todo en las secciones de comentarios de periódicos y blogs. Allí se juntaban contestatarios disparatados y conspiracionistas, personas sin ningún contexto personal compartido más allá del deseo de registrar su objeción a la cobertura periodística sobre el clima. Aquello creaba una sensación de objetivo común. Y la ubicación de los comentarios de los periódicos — justo debajo del artículo— les daba una visibilidad inusual: eso hacía que los lectores diarios tuvieran una falsa impresión sobre la popularidad de esas personas, con lo cual allí pescaban grandes cantidades de nuevos adeptos.

¿Podía ser que YouTube estuviera haciendo algo parecido?, se preguntaba Kaiser. Una de las redes que la plataforma había entrelazado con el Gamergate era la derecha alternativa. La extrema derecha llevaba décadas ahí, eso lo sabía. Pero en internet, ahora parecía entretejida con círculos de las redes sociales que poco tenían que ver con la política y se fusionaba para formar algo más grande, algo nuevo. Una vez terminado el doctorado, Kaiser se puso en contacto con el suizo Adrian Rauchfleisch, un estudiante de posgrado con unos intereses similares y con talento para la programación, quien se convertiría en su compañero de investigación durante largo tiempo. Los dos jóvenes readaptaron las herramientas que Kaiser había desarrollado para monitorizar el escepticismo ante el cambio climático, ahora con el objetivo de entender el ascenso de la extrema derecha en Alemania.

Sin embargo, sospechaban que todas las lecciones aprendidas se aplicarían también a Estados Unidos. Era verano de 2016, cuando Donald Trump iba subiendo en las encuestas. Kaiser identificó los vínculos de Trump con nacionalistas blancos alineados con el Gamergate que había visto en YouTube, además de los paralelismos con la derecha alternativa de su país. Aunque el conjunto de datos inicial de los dos investigadores sería alemán, sus preocupaciones se centraron cada vez más en Estados Unidos.

Empezaron con YouTube, donde habían iniciado su andadura varias de las voces principales de la derecha alternativa, tanto alemanas como estadounidenses.⁵Muchos calificaban la plataforma como su base digital de operaciones. No obstante, era la menos

estudiada de las grandes plataformas, seguía siendo en buena medida una caja negra. Kaiser y Rauchfleisch cartografiarían YouTube y trazarían cómo las recomendaciones de la plataforma guiaban a los usuarios. Sería lo que Serrato había hecho con los vídeos de Chemnitz pero a una escala miles de veces mayor; quizás millones.

Programaron un ordenador para que rastrease las recomendaciones de YouTube de varios canales alemanes muy conocidos, luego las siguientes, y así sucesivamente.⁶ Repitieron el proceso una y otra vez buscando patrones en las elecciones del sistema. Puesto que las recomendaciones de vídeos pueden ser caóticas —varían día a día a medida que el algoritmo aprende y persona a persona basándose en los perfiles individuales de cada usuario—, los dos investigadores también rastrearon una de las pocas recomendaciones de YouTube estandarizadas: una lista de «canales relacionados» que el algoritmo generaba en la parte inferior de cada canal. Eso proporcionaba otra capa de información sobre el pensamiento del algoritmo, lo cual ayudaba a Kaiser y Rauchfleisch a separar los patrones del ruido. Era un poco una forma de limpiar la casa en términos metodológicos —una comprobación para asegurarse de que obtenían resultados rigurosos— que más adelante, cuando sus resultados provocaron indignación en altos estratos del Gobierno estadounidense, se convertiría en una fuente de conflicto con YouTube.

Los investigadores pensaban que los resultados se parecerían a una nube: miles de canales de contenido variado bastante desorganizados. Por el contrario, la red presentaba una nítida serie de agrupaciones, organizadas una al lado de otra, como el mapa de un metro. Estaban asombrados. Porque el hecho de que los sistemas de YouTube analizaran y clasificaran miles de millones de horas de vídeo en tiempo real y luego dirigieran a miles de millones de usuarios por la red con ese nivel de precisión y uniformidad era un hito tecnológico increíble, que demostraba la sofisticación y la potencia del algoritmo.

Los vídeos sobre política estaban aún más concentrados. El sistema había creado una superagrupación —parecía una borrasca en un mapa del tiempo— a partir de varias comunidades que Kaiser y Rauchfleisch esperaban ver separadas. Medios de comunicación alternativos, comentarios de centroderecha, extremistas de ultraderecha, universitarios neorracistas y teóricos de la conspiración

con sombreros de papel de aluminio estaban todos conectados. Con el objetivo de comprobar si esa superagrupación afectaba al comportamiento de los usuarios, los investigadores rastrearon todos los comentarios sobre los vídeos durante un periodo de tres años, dos millones en total, y luego indagaron en la actividad de los comentaristas (de forma anónima para garantizar la privacidad) por toda la plataforma.⁷

Los resultados fueron como se habían temido. Usuarios que empezaban comentando solo un subgrupo de vídeos —por ejemplo, canales de noticias de centroderecha— terminaban escribiendo comentarios en otros canales de la superagrupación. «Con el tiempo, la red se volvía más densa y compacta», dijo Kaiser. A medida que a los usuarios se les dirigía por medio del algoritmo entre comentaristas de la derecha mayoritaria, conspiracionistas estrafalarios y racistas de sótano, empezaban a tratar esas páginas, antes alejadas, como parte de una comunidad unificada. Su comunidad. YouTube, dijo Kaiser, había creado una nueva «identidad colectiva».

El proyecto los ocupó durante un año, hasta verano de 2017. La victoria de Trump había agudizado la situación. Sabían que, si estaba ocurriendo algo parecido en Estados Unidos, eso significaba que las fuerzas que habían identificado en Alemania ahora estaban dominando el país más poderoso del planeta. Así pues, entender ese movimiento —sus componentes y qué vinculación había entre ellos— era una cuestión de suma urgencia. Antes de que tuvieran la oportunidad de analizarlo, algo que se parecía un montón a la comunidad formada en el mundo digital que habían encontrado en Alemania salió a la superficie en lo que fuera el corazón de la Confederación.

En agosto de 2017, varios centenares de manifestantes se dieron cita en Charlottesville, en Virginia, para lo que externamente llamaron una concentración contra el plan de la ciudad de eliminar una estatua del general confederado Robert E. Lee. En las redes sociales donde se organizaron, llamaron a su acto «Unir a la derecha». Su alianza tenía algo de particular. Grupos radicales con poca o ninguna vinculación hasta entonces se habían unido de repente. Llegaron seguidores en unas cantidades antes inimaginables, como si hubiesen aparecido de la nada. Muchos no formaban parte de ningún grupo, eran agentes libres que participaban porque sí.

Su ímpetu se convirtió en una demostración de fuerza. Cientos de personas portaban antorchas y ondeaban banderas nazis y confederadas mientras coreaban «Los judíos no nos van a sustituir». Con cascos en la cabeza y porras en las manos, muchos se enfrentaron a la policía. En el segundo día, poco después de que el gobernador declarase el estado de emergencia, uno de los manifestantes embistió adrede a una multitud de manifestantes contrarios con su Dodge Challenger, con lo cual hirió a decenas de personas y mató a una mujer de treinta y dos años llamada Heather Heyer.

No era difícil entender cómo la extrema derecha se había envalentonado. Trump había utilizado la presidencia para defender el movimiento y muchas de sus causas y conspiraciones. Sin embargo, eso no explicaba cómo esos grupos se habían unido viniendo de entornos tan distintos, cómo esa unificación había sido tan transversal y tan rápida. Resultó que la concentración y el grupo que había detrás se habían constituido en las redes sociales. Pero incluso el acto, a pesar de su terrible importancia, apuntaba a algo mucho más amplio, y en última instancia más desestabilizante, que estaba ocurriendo en internet.

La derecha alternativa digital, aún cercana a sus raíces del Gamergate, estaba centralizada en una abarrotada subsección de Reddit llamada The_Donald.⁸ Aunque en su mayor parte estaba gobernada por su propia cultura del troleo, tomaba su inspiración de extremistas como Andrew Anglin, la estrella de 4chan convertida en neonazi. Anglin había declarado un «verano del odio» y animaba a organizar manifestaciones en el mundo real.⁹ A medida que aumentaba en internet el sonido de tambores para la organización de un megaevento, un activista llamado Richard Spencer empezó a organizar uno en Charlottesville. Spencer, quien había acuñado el término *alt-right* ('derecha alternativa') como una versión amistosa para referirse al nacionalismo blanco y así atraer a chavales universitarios, vivía en las redes sociales y sobre todo era un invitado habitual en canales de derechas de YouTube. Una vez le dijo a un periodista: «Dimos a luz a la derecha alternativa a base de memes».¹⁰ Mientras Spencer y Anglin presionaban para que se organizase la manifestación, los moderadores del subreddit The_Donald apoyaban el evento y animaban a participar en él.¹¹

Pero fue en Facebook donde el evento pasó de un encuentro de

miembros de Reddit a convertirse en la presentación en sociedad del extremismo transversal, determinó Megan Squire, experta en el extremismo digital de la Universidad de Elon. En un estudio que analizaba la importancia de la plataforma, Squire identificó 1.870 grupos y eventos de ultraderecha de Facebook.¹² Los encontró revisando los perfiles de conocidos extremistas, buscando términos asociados con la extrema derecha y, claro está, siguiendo «grupos recomendados» por el algoritmo de Facebook, que la guió fácilmente hacia un universo de odio. Rastreó los miembros de cada grupo. Luego hizo un análisis de la red, como había hecho Kaiser con YouTube, para visualizarlo todo. Cada grupo de Facebook era un nodo. Cuantos más miembros compartían dos grupos, más cerca estaban esos dos nodos.

En una red social más neutra, eso podría haber resultado en cinco o seis agrupaciones distintas —por ejemplo, nostálgicos de la Confederación, neonazis, milicias antigobierno, círculos de memes de la derecha alternativa— que no se mezclasen en el mundo físico. Pero en Facebook, al igual que en YouTube en Alemania, la plataforma fusionaba esas comunidades que de otro modo estarían separadas y creaba algo nuevo por completo. Y en el centro de todo eso: la página del evento «Unir a la derecha».

Jonas Kaiser, que ahora estaba en Estados Unidos tras aceptar un trabajo temporal como profesor en Harvard, estaba horrorizado. «Siendo alemán, probablemente era un poco ingenuo», dijo. En Alemania, en las concentraciones de la extrema derecha, me contó, «salen a la calle algunos nazis y gritan, pero por lo general hay muchos más manifestantes contrarios que nazis». La concentración nacionalista blanca de Charlottesville había sido algo mucho más radical. «Aquello hizo aumentar la urgencia —dijo—, tanto para mí como para Rauchfleisch.»

Entonces comenzaron a aplicar las técnicas que habían desarrollado en Alemania, esta vez al YouTube anglófono. En ese momento, estaban aprendiendo cómo identificar mucho más que simples conexiones. Podían cartografiar no solo la red en conjunto, como habían hecho en Alemania, sino cómo el algoritmo guiaba a los usuarios dentro de ella, qué canales el sistema trataba como puntos de entrada y cuáles como destinos. Era la diferencia entre hacer el plano de una ciudad o el mapa de su tráfico.

En enero de 2018, Kaiser ya había acumulado suficientes pruebas

para disponerse, poco a poco, a hacer públicas sus conclusiones. Contó en un seminario de Harvard que empezaba a creer que la integración de la extrema derecha, de la que formaba parte la concentración de Charlottesville, «no la habían hecho los usuarios», por lo menos no del todo, sino que en parte «se había creado mediante el algoritmo de YouTube». ¹³ Sabía que él y Rauchfleisch estaban cerca de demostrarlo. Pero ninguno de los dos podía haber anticipado lo que encontrarían a lo largo del camino.

CRISIS Y SOLUCIÓN

En los rincones más oscuros del mundo digital de la extrema derecha, la influencia de YouTube ya era una cuestión aceptada de manera general. En marzo de 2018, un usuario de un destacado foro neonazi inició un hilo de debate con el título «¿QUÉ TE TRAJO AL MOVIMIENTO?». ¹⁴ El foro, llamado «the Right Stuff» ('lo que hay que tener'), había crecido a lo largo de la década de 2010 y había dado lugar a muchos de los elementos distintivos de la derecha alternativa. A su creador, Mike Peinovich, se le considera el inventor del «(((eco)))», ¹⁵ en el que los usuarios ponen los nombres de supuestos judíos entre paréntesis triples. El meme, iniciado en el pódcast *Daily Shoah*, llamado así en tono de burla por la denominación del Holocausto en hebreo, se volvió omnipresente en Twitter y Facebook. En panfletos que anunciaban la concentración «Unir a la derecha», el nombre de Peinovich, como muestra de su popularidad, aparecía en segundo lugar después del de Richard Spencer. ¹⁶

Muchos usuarios respondieron al hilo presentando sus historias. Uno tras otro, contaban que habían adoptado sus opiniones de forma gradual, siempre en las redes sociales, a menudo con la ayuda de los algoritmos. «Yo formaba parte de los anti-SJW», ¹⁷ escribió un usuario, en referencia a los «social justice warriors» ('luchadores por la justicia social'), vilipendiados de forma habitual en 4chan y Reddit. Y añadía, empleando el término del argot de internet *based*, un adjetivo para denominar algo transgresor: «Y ahora estoy aquí. Gracias algoritmo que me recomendó vídeos del YouTube “transgresor”». Como a muchos, las recomendaciones de YouTube lo habían llevado, decía, primero a voces ultraconservadoras. Luego a nacionalistas blancos,

después a supremacistas blancos, más adelante a neonazis supuestamente irónicos y, al final, a neonazis auténticos.

También el propio Peinovich había seguido una ruta parecida. Programador web de Nueva York, moderado en política y con una esposa judía, Peinovich había desarrollado problemas de salud a finales de la década del año 2000. Encerrado en casa, se pasó «horas en foros de debate político de Facebook y Reddit, donde dio rienda suelta a su lado contestatario»,¹⁸ explicaron unos amigos suyos al *New Yorker*. Después de que el tiempo y los algoritmos hicieran su trabajo, Peinovich se había convertido en un auténtico nazi convencido.

Otro tema recurrente surgió de los relatos de los usuarios de «the Right Stuff». Pocas veces lo que atraía a la gente era el odio. Más bien era contenido que apelaba a sentimientos de alienación o falta de sentido: la misma anomia que había unido a Adam y a otras personas obsesionadas con 4chan. «Vi un vídeo de Millennial Woes sobre la depresión (estaba fatal)», decía uno de los usuarios al describir el primer paso de su viaje. Millennial Woes, el brazo en YouTube del supremacista blanco escocés Colin Robertson, publicaba vídeos prometiendo desvelar la verdadera causa de la infelicidad de los jóvenes: la sociedad, decía, pretendía desbancar a los hombres blancos del lugar natural en lo alto de la jerarquía social. «Mencionaba varias cosas de las que nunca había oído hablar (neorreacción, derecha alternativa) —seguía contando el usuario de «the Right Stuff»—. Encontré a Richard Spencer. Encontré TRS. Mi vida cambió por completo.»

Una de las puertas de entrada digitales más importantes a la derecha alternativa es la página de YouTube de Jordan Peterson, un profesor de Psicología canadiense.¹⁹ En 2013, Peterson empezó a subir vídeos en los que abordaba, en medio de una esotérica filosofía jungiana, el malestar de los hombres jóvenes. Ofrecía consejos de vida (limpia tu habitación, siéntate erguido) junto con llamamientos contra la igualdad racial y de género porque ponían en peligro «el espíritu masculino».²⁰

Buscar en YouTube «depresión» o ciertas palabras clave relacionadas con la autoayuda a menudo conduce a Peterson.²¹ La inusual duración de sus vídeos, de sesenta minutos o más, va de la mano con el impulso del algoritmo a maximizar el tiempo de visualización. Como también su método, al estilo de los currículos

universitarios, de serializar sus argumentos a lo largo de semanas, lo cual requiere volver para la siguiente clase y la otra. Pero, por encima de todo, Peterson apela a lo que el sociólogo Michael Kimmel denomina «privilegio ofendido».²² Durante generaciones, los hombres blancos esperaban y recibían un trato prioritario y gozaban de un estatus especial. A medida que la sociedad fue avanzando poco a poco hacia la igualdad, esas ventajas, aunque siguieron siendo sustanciales, disminuyeron. Algunos hombres blancos se adaptaron. Otros se rebelaron. Otros solo sabían que tenían la sensación de que les habían arrebatado algo. Peterson y otros personajes les dan una vía para explicar esos sentimientos de injusticia —las feministas y la gente de izquierdas están destruyendo el espíritu masculino— y una lista de respuestas fáciles. Limpia tu habitación. Siéntate recto. Reafirma las jerarquías tradicionales.

«Las personas que forman esas comunidades a menudo están ahogándose en la desesperación y son susceptibles de ser corrompidas por personajes lúcidos y carismáticos que les ofrecen lo que parece el único salvavidas»,²³ tuiteó un antiguo adepto. El algoritmo de YouTube, en muchos casos, aprovechaba ese descontento y recomendaba canales que llevaban el mensaje de Peterson hasta el extremo. «Los estadios típicos —escribió el usuario, eran— Jordan Peterson -> Stefan Molyneux -> Millennial Woes.» (Molyneux, un supremacista blanco que se presenta como un terapeuta y un filósofo «que solo hace preguntas», había trabajado duro en secreto antes de unirse a YouTube. Cuando la presión pública obligó a la compañía a eliminar su página en junio de 2020, tenía más de 900.000 suscriptores y llegaba a muchos más usuarios mediante las recomendaciones automatizadas.) Era justo el tipo de recorrido algorítmico que Kaiser estaba cartografiando en su laboratorio, el mismo que descubrió estudiando las rutas de YouTube una y otra vez.

Los datos dan a entender que esa secuencia promocional está convirtiendo usuarios a gran escala. Los usuarios que comentan los vídeos de Peterson tienen el doble de probabilidad de aparecer después en los comentarios de canales de YouTube de extrema derecha, según descubrió un estudio de Princeton.²⁴ El propio Peterson no recomienda los canales: es el algoritmo quien hace la conexión. Esa era la otra pieza esencial del rompecabezas que Kaiser y Rauchfleisch estaban trabajando para resolver. Estaban midiendo cómo YouTube

movía a sus usuarios. Pero sus planos de la red no sabían decirles por qué el sistema tomaba las decisiones que tomaba. Hicieron falta psicólogos e investigadores del extremismo para hallar la respuesta.

Las plataformas sociales habían llegado, aunque fuera sin querer, a una estrategia de captación adoptada por generaciones de extremistas. El académico J. M. Berger lo denomina «el constructo crisis-solución».²⁵ Cuando una persona se siente desestabilizada, a menudo busca una identidad de grupo fuerte para recuperar la sensación de control. Esa identidad puede abarcar tanto como la nacionalidad o tan poco como una parroquia. Las identidades que prometen recontextualizar las adversidades individuales en un conflicto más amplio tienen un especial atractivo. No eres infeliz a raíz de tu batalla para luchar contra unas circunstancias personales; eres infeliz por culpa de Ellos y porque Ellos Nos persiguen. Eso hace que esas adversidades sean comprensibles y, como uno ya no las está afrontando solo, dan mucho menos miedo.

Crisis-solución: hay una crisis, el exogrupo es responsable, el endogrupo proporciona la solución. Si esa sensación de conflicto aumenta demasiado, puede llegar al extremo de la radicalización, en que uno ve al exogrupo como una amenaza inmutable ante la cual solo es aceptable la victoria total y absoluta. «El tamaño de la crisis se vuelve más extremo y la solución prescrita se hace más violenta», escribió Berger, hasta que destruir al exogrupo se convierte en el núcleo de la identidad compartida del endogrupo. «La actual generación de plataformas de redes sociales —añadía Berger— acelera la polarización y el extremismo de una minoría significativa»,²⁶ lo que hace posible y alienta justo ese ciclo.

El número de víctimas ya estaba subiendo. En 2014, en una población cercana a Santa Bárbara, un chico de veintidós años que había dejado la universidad llamado Elliot Rodger apuñaló mortalmente a tres personas en su bloque de pisos, luego subió un vídeo a YouTube anunciando su «venganza contra la humanidad» por verse «forzado a llevar una existencia de soledad, rechazo y deseos insatisfechos porque las chicas nunca se han sentido atraídas por mí». ²⁷ Entonces Rodger condujo hasta una residencia estudiantil para chicas, llamó a la puerta y disparó contra tres mujeres, dos de las cuales murieron. Regresó a su coche y recorrió las calles de la zona mientras disparaba y arrollaba a peatones, muchos de ellos mujeres, y

terminó suicidándose.

Que se sentía solo, sí; pero solo no estaba: durante meses había estado maquinando en foros virtuales, luego en YouTube, donde sus vídeos alcanzaron tales niveles de odio que una vez sus propios padres llamaron a la policía. Rodger era producto de una comunidad digital, que se extendía por 4chan, Reddit y YouTube, cuyos miembros se llaman a sí mismos «*incels*» en referencia a su celibato involuntario (*involuntary celibacy*, en inglés).²⁸

Los foros de *incels* habían empezado siendo lugares para compartir relatos sobre el sentimiento de soledad. Los usuarios hablaban de cómo sobrellevar una vida sin abrazos. Pero las normas de las redes sociales a la hora de dar prioridad, de acaparar la atención, seguían prevaleciendo. Las voces más vehementes ganaron peso. Las opiniones se extremaron.²⁹ La indignación predominante en las plataformas reinterpretó malestares individuales como una lucha tribal entre un Nosotros y un Ellos. En internet, los *incels* adoptaron una creencia nuclear radicalizadora: las feministas están conspirando para subyugar y emascular nuestra clase marginal de hombres con bajo estatus. Endogrupo y exogrupo. Crisis y solución.

En 2021, ya se habían producido cincuenta asesinatos reivindicados por chicos que se autoproclamaban *incels*, una oleada de violencia terrorista.³⁰ Un *youtuber* de poca monta mató a tiros a cuatro mujeres en un centro de yoga de Florida. Un superusuario de Reddit arrolló con su furgoneta a una multitud de peatones en Toronto. Un *íncel* de 4chan publicó en directo mientras asesinaba a una muchacha de diecisiete años, y fue aplaudido por otros usuarios. Luego intentó suicidarse. Rodger sigue siendo muy idolatrado como un héroe entre los *incels*.³¹

El movimiento estaba en los márgenes de los márgenes, eclipsado por el Pizzagate o la derecha alternativa. Pero daba pistas del potencial de las redes sociales para galvanizar la anomia de los hombres jóvenes blancos y formar comunidades de extremismo: un fenómeno cada vez más extendido. Cuando un grupo paramilitar de extrema derecha llamado «the Oath Keepers» encuestó a sus 25.000 miembros sobre cómo habían llegado al movimiento, la respuesta más frecuente fue Facebook, seguido de YouTube.³²

YouTube puede ser una herramienta adoctrinadora muy efectiva porque mueve a los usuarios por fases. Jordan Peterson cuenta a los

usuarios que sus problemas personales derivan de un conflicto que los enfrenta a los luchadores por la justicia social: crisis. Millennial Woes los une para defenderse de forma colectiva contra las feministas y las minorías que se oponen a ellos: solución. Canales más extremos suben la apuesta de esa guerra hasta el genocidio blanco o la sumisión judía y animan de un modo implícito a enfrentar la amenaza como sea necesario.

YouTube no hace esto en todos los casos. A los consumidores de Jordan Peterson también se les redirige, por ejemplo, a académicos populares o vídeos de autoayuda. Pero las plataformas suelen privilegiar las conexiones radicalizadoras por un motivo: funciona. A los extremistas les gusta la secuencia crisis-solución porque prepara a la gente para actuar. A los algoritmos les gusta porque capta la atención y la pasión de los usuarios, lo que convierte la navegación por internet en una cuestión de identidad, comunidad e incluso fanatismo; y, por tanto, en más tiempo de visualización.

YouTube actualizó sus algoritmos durante 2016 y 2017 añadiéndoles un sistema llamado Reinforce,³³ que recomendaba a los usuarios subgéneros que antes desconocían. Aunque nunca hubieras buscado vídeos de entrada a la derecha alternativa del estilo de Peterson, podía ser que te arrastrasen hacia ellos de todos modos, solo para ver si funcionaba. De repente, había historias de radicalización en YouTube por todas partes, con unos detalles que se repetían con una regularidad maquinal. «A uno de mis amigos más cercanos lo radicalizó YouTube —tuiteó Chris Sacca, inversor de Silicon Valley y exalumno de Google—. Todo empezó hace algunos años con vídeos “que daban que pensar” y “contestatarios”. Pero gracias al algoritmo de los vídeos recomendados, se volvió más misterioso y más violento, perdió a su mujer, a sus hijos y a sus amigos, y ninguno de nosotros sabe dónde está a día de hoy.»³⁴

David Sherratt, un antiguo extremista, contó al *Daily Beast* que su descenso había empezado, a los quince años, con vídeos de videojuegos.³⁵ El sistema le recomendó vídeos proateísmo, que cuentan a chavales de ciencias que forman parte de una minoría ultrarracional asediada por los luchadores de la justicia social. Luego vídeos antifeministas, luego vídeos sobre los «derechos de los hombres» vinculados al mundo íncel —algunos de los cuales publicó él mismo—, luego vídeos directamente neonazis.

En 2018, el medio de comunicación Bellingcat exploró un archivo de salas de chat privadas de extrema derecha que contenían un total de cientos de miles de mensajes.³⁶ Los investigadores buscaron fragmentos en los que los usuarios hubiesen mencionado cómo habían llegado a la causa. El elemento más común que citaban: YouTube. Empezaban con vídeos banales, decían muchos, luego les aparecían recomendados canales que eran cada vez más extremos. Todo el mundo contaba la misma historia.

«Los algoritmos de YouTube me condujeron por un recorrido de criminales del pensamiento todos desacomplejados»,³⁷ escribió la novelista Meghan Daum sobre su deriva hacia un movimiento que está a medio camino de la derecha alternativa, sobre todo en YouTube, que se autodenomina «*dark web* intelectual». Jordan Peterson es un miembro destacado. A medida que la plataforma la arrastraba de unos vídeos que cuestionaban el feminismo a otros que insinuaban que los cerebros de las mujeres estaban configurados para hacer las tareas del hogar, de llamar a los negros racistas a rechazar el pluralismo racial, ver YouTube se convirtió en «lo que hacía en lugar de ver la tele (y, en muchos casos, en lugar de leer libros, escuchar música o limpiar el piso)». Era algo más que entretenimiento. Esos canales eran sus «amigos de YouTube», bálsamo para un matrimonio fracasado y sus sentimientos de aislamiento. Eran comunidad. Eran identidad.

Había caído en lo que se conoce como una madriguera. El término antes describía cualquier noche o tarde pasada siguiendo las recomendaciones de YouTube hasta donde te llevaran. Quizás empezabas viendo un vídeo de humor de tus favoritos, luego te recostabas mientras el sistema te guiaba por los éxitos del humorista. Pero, en 2018, después de implementarse el sistema Reinforce, el término «madriguera» se refirió cada vez más a seguir canales de política en YouTube que conducían al extremismo. Los usuarios se caían por esos agujeros tanto si buscaban vídeos políticos como si no, y a menudo terminaban en lugares a los que no pretendían ir: lugares aún más perturbadores que la extrema derecha.

EL PROBLEMA DE ALEX JONES

Durante la primavera de 2018, Kaiser y Rauchfleisch hicieron test

automatizados de las recomendaciones de YouTube, con lo cual cartografiaron cómo la plataforma interconectaba los canales y qué comunidades construía.³⁸ En abril juntaron los resultados. Descubrieron que el sistema de YouTube hacía tres cosas insólitamente bien.

Primero, entremezclaba agrupaciones de canales del todo originales. Una de esas agrupaciones mezclaba médicos que daban consejos sanitarios, profesores de yoga que defendían ideas sobre la alineación de los chakras, charlatanes que vendían remedios mágicos y radicales antivacunas. A un usuario que empezaba por alguno de esos temas se le conducía por los demás. No había nada inherente que los conectase más allá de la conclusión alcanzada por la IA de que mostrarlos uno al lado de otro haría que el usuario pasase más tiempo en la plataforma. Aunque los otros dos descubrimientos recibirían la mayor atención, Kaiser insistía todo el tiempo en la importancia del primero. Al igual que Facebook o Twitter, YouTube es sobre todo una experiencia social. Los usuarios comentan, ponen «me gusta» y comparten, todos ellos participantes de una actividad comunitaria. A los creadores de vídeos se les anima a dirigirse a los seguidores, a pedir la opinión y los deseos de los usuarios y a participar ellos mismos en los comentarios. Así que, cuando YouTube agrupaba a varios miles (o millones) de usuarios en varias decenas (o cientos) de canales, aquello se convertía en una auténtica comunidad. Se convertía, como había descrito Meghan Daum con tanta gratitud, en una identidad.

Eso hacía que sus otros descubrimientos fueran tan importantes. El segundo hallazgo era que por lo general —como se había sospechado durante mucho tiempo— las recomendaciones de YouTube arrastraban hacia los rincones más extremos de cualquier red en la que se encontrase el usuario. Si alguien veía un vídeo de la CNN o Fox News, la probabilidad de que le recomendaran un vídeo conspirativo era lo bastante elevada para que, con el tiempo, esa fuera la dirección a la que el usuario medio tendía a ir.

Y, por último, el tercer descubrimiento. En una versión aún más peligrosa de lo que habían visto en Alemania, las recomendaciones del sistema agrupaban canales de derecha mayoritarios, e incluso algunos canales de noticias, con muchos de los incitadores de odio, íncels y teóricos de la conspiración más vehementes de la plataforma. A

medida que el algoritmo arrastraba a los usuarios hacia esas voces extremas, concedía a dichos canales una influencia desmedida sobre los relatos, los programas políticos y los valores para la mayoría.

Eso iba más allá de ampliar el alcance de la extrema derecha. Significaba unir a una comunidad más amplia alrededor de ellos. Y a una escala —de millones de personas— que los organizadores del acto de Charlottesville solo podían concebir en sueños. Aquí, por fin, había una respuesta a por qué había habido tantas historias de personas que caían por madrigueras de la extrema derecha. Alguien que llegaba a YouTube interesado en temas vinculados con la derecha, como las armas o la corrección política, terminaba redirigido a un mundo construido por YouTube de nacionalismo blanco, violencia misógina y un conspiracionismo enloquecido, y luego era arrastrado hacia los extremos.

Había un canal que ocupaba de forma llamativa el centro de la red, un agujero negro hacia el cual atraía la gravedad algorítmica de YouTube: Alex Jones. Locutor de radio provocador desde los años noventa, Jones llevaba tiempo cultivando un público de cascarrabias, camioneros nocturnos y personas aburridas. De carácter animado y con voz cavernosa, se regodeaba en conspiraciones elaboradas no tanto para convencer a sus oyentes como para retener su atención: el Gobierno fue el responsable de los atentados de Oklahoma City; el Gobierno echó sustancias químicas al agua que llega a las casas para convertir a los ciudadanos en homosexuales. Montaba escenas para captar la atención, como hacer callar a gritos al entonces gobernador George W. Bush mientras daba un discurso para exigir la abolición de la Reserva Federal. Vendía suplementos nutricionales falsos y kits de supervivencia, con lo cual amasó una fortuna. Luego abrió un canal de YouTube y se convirtió, al instante, en una figura influyente. Sus declaraciones llegaban a Fox News y blogs de derecha. Su nombre se invocaba en mítines conservadores, actos políticos de base, incluso entre altos cargos de Washington, aunque fuera solo para reconocer su nueva capacidad de influencia.

¿Cómo había ocurrido eso? Sin duda, los estadounidenses se habían decantado hacia la dirección de Jones. Y él había usado internet para eludir a los organismos de supervisión de los medios que lo mantenían en buena medida alejado de la televisión. Pero Kaiser tenía pruebas de otro factor. Sus datos apuntaban a que el algoritmo

de YouTube recomendaba los vídeos de Jones con mayor frecuencia que aquellos de casi cualquier otro canal de noticias o política. Los promovía con una especial vehemencia junto con vídeos de derechas, con lo cual introducía a Jones en la causa. A menudo usaba sus vídeos como puerta de entrada al extremismo: a un usuario que veía sobre todo conspiraciones le podía aparecer una diatriba de Alex Jones como punto de entrada a contenido nacionalista blanco o viceversa.

En abril de 2018, tan solo semanas después de que la violencia en Sri Lanka hubiera llamado la atención sobre la capacidad de Facebook de causar estragos, Kaiser publicó sus resultados sobre la creación sistemática por parte de YouTube de redes de extrema derecha,³⁹ junto con un artículo en *BuzzFeed News* que publicitaba el estudio, en un ensayo explicando que YouTube «empuja muchos canales hacia el centro gravitatorio de una burbuja de derechas más grande».⁴⁰ Él y Rauchfleisch alertaban de que «ser conservador en YouTube significa que uno está a tan solo uno o dos clics de canales de extrema derecha, teorías de la conspiración y contenido radicalizador». Ahí, al fin, había pruebas de que YouTube se había convertido en una fuerza que favorecía la radicalización a gran escala.

Otros pronto confirmaron el «conducto de radicalización», como lo denominó el investigador brasileño Manoel Horta Ribeiro.⁴¹ Su equipo, tras analizar 72 millones de comentarios de 330.000 vídeos, halló que «los usuarios migran de forma sistemática de contenidos más leves a más extremos». Los usuarios de derechas, un grupo enorme, pasaban de contestatarios de la «*dark web* intelectual» como Jordan Peterson a voces de la derecha alternativa como Milo Yiannopoulos o a adalides del odio como los neonazis Andrew Anglin y Mike Peinovich. Y los usuarios se desplazaban en paralelo a las recomendaciones de YouTube, otra prueba de que era el algoritmo lo que los guiaba.

Esa primavera, tras un tiroteo en un colegio, la destacada página de «tendencias» de YouTube empezó a dar visibilidad a un vídeo de Alex Jones en que se afirmaba que la violencia se había simulado. Jones venía difundiendo versiones de esa idea desde la masacre de Sandy Hook, en 2012, cuando había calificado a los veinte niños y seis maestras asesinados de «actores que hacían de víctimas» en un vago complot del Gobierno para justificar la confiscación de armas o la imposición de la ley marcial. La conspiración se había propagado por

YouTube desde entonces, consumida por un número cada vez mayor de usuarios que, indignados, organizaron campañas de acoso durante años contra las familias de los niños asesinados. Algunos padres se escondieron, y varios de ellos presentaron tres demandas contra Jones por difamación. (En 2021, Jones perdió las tres.) Mientras tanto, YouTube siguió promoviendo los vídeos, que en 2018 tenían 50 millones de visualizaciones en conjunto.⁴² Cuando el sistema de YouTube publicó el vídeo más reciente de Jones en su página de tendencias, aquello indicó con una claridad insólita el entusiasmo con que los algoritmos de la plataforma lo promovían.

Miembros del equipo de políticas de YouTube recomendaron modificar el algoritmo de la página de tendencias para impedir que enlazase a Jones y a otras fuentes sin credibilidad.⁴³ La recomendación les fue denegada. Aunque su presencia más destacada era en YouTube, donde tenía miles de millones de visualizaciones, Jones también llegaba a millones de seguidores en Facebook y Twitter, y esas compañías también recibieron presiones para suprimir los megáfonos que habían puesto a su disposición.

Los padres de Noah Panzer, un niño de seis años asesinado en Sandy Hook, publicaron una carta abierta a Mark Zuckerberg. Vivían escondidos —escribían—, atormentados por años de acoso y amenazas de muerte que tenían su origen en destacados grupos de Facebook sobre conspiraciones. Los padres de Sandy Hook —decían— habían librado «una batalla casi inimaginable con Facebook para ofrecernos la más básica de las protecciones» contra la incitación que emanaba de la plataforma. «Nuestras familias están en peligro —escribían— como consecuencia directa de los cientos de miles de personas que ven y creen las mentiras y los discursos de odio, que vosotros habéis decidido que hay que proteger.»⁴⁴

Pero las compañías no cedieron. Al preguntarle por qué Facebook no bloqueaba a Alex Jones, el vicepresidente de la empresa responsable del canal de noticias, John Hegeman, dijo: «Creamos Facebook para que fuese un lugar donde pudieran tener voz personas diversas». ⁴⁵ Zuckerberg, ante la misma pregunta, improvisó hablando sobre la naturaleza de la libertad de expresión: «Soy judío, y hay una serie de personas que niegan que el Holocausto tuviese lugar. A mí eso me parece muy ofensivo. Pero, al fin y al cabo, no creo que nuestra plataforma deba bloquearlas, porque pienso que hay cosas en las que

personas distintas se equivocan. No creo que se equivoquen adrede».⁴⁶

Aquello era Silicon Valley en su versión clásica. Si Zuckerberg estaba dispuesto a sacrificar el consenso histórico sobre el intento de exterminio de sus antepasados a favor de un ideal tecnolibertarista de libertad de expresión, entonces todo el mundo debía hacer lo mismo. Y, al igual que muchos de los directivos de Silicon Valley, él parecía estar viviendo en un universo paralelo donde las plataformas eran unos recipientes neutros que no participaban en la conformación de las experiencias de los usuarios, donde la única consecuencia en el mundo real era que alguien podía ofenderse, y donde la sociedad apreciaría la sensatez de permitir que los negacionistas del Holocausto tuvieran voz.

Ese era el problema de Alex Jones: Silicon Valley no podía forzarse a actuar a favor del interés público, ni siquiera en su propio interés, extirpando un problema que encarnaba, personificaba incluso, las creencias de su sector, inscritas en los sistemas que, a fin de cuentas, habían elevado a Jones hasta ser lo que era. Participación de los usuarios equivale a valor. Una mayor interconexión hace aumentar el entendimiento. La libertad de expresión vence a la intolerancia. Actuar significaría admitir que esos ideales eran erróneos, peligrosos. Si se tiraba de ese hilo, podía desmoronarse todo.

Las críticas empeoraron. Ese julio, Facebook suspendió a Jones durante tres meses, una concesión a quienes lo criticaban. Al final, en agosto, Apple retiró varios de los programas de Jones de su servicio de pódcast. Facebook y YouTube hicieron lo propio al cabo de unas horas y también lo prohibieron. Solo Twitter se mantuvo desafiante. «Vamos a aplicar a Jones los mismos estándares a los que sometemos todas las cuentas, sin tomar medidas excepcionales para sentirnos bien a corto plazo»,⁴⁷ tuiteó el director ejecutivo de Twitter, Jack Dorsey. Pero Twitter también acabó eliminando el perfil de Jones. Aunque las compañías habían cedido ante la presión social, no parecían aceptar que la sociedad tenía razón. Zuckerberg dijo más adelante al *New Yorker*: «No creo que sea lo correcto expulsar a una persona por decir algo que no se ajusta a la verdad de los hechos».⁴⁸

En Twitter, Dorsey, a pesar de su actitud desafiante, estaba virando hacia los cambios profundos, en la naturaleza central de la plataforma, a los que Silicon Valley se había resistido durante mucho tiempo. O, por lo menos, parecía estar haciéndolo. La capitalización

de mercado de Twitter permanecía por debajo de su máximo, en abril de 2015, desde antes de que Dorsey había vuelto a tomar las riendas, y su base de usuarios apenas había aumentado en el año anterior. Pero, en lugar de potenciar sus algoritmos o remodelar la plataforma para dar prioridad a las disputas y las emociones, como habían hecho YouTube y Facebook ante el estancamiento de los indicadores, Dorsey manifestó que toda la idea que había detrás de las redes sociales era tóxica.

«No previmos ni entendimos por completo las consecuencias negativas en el mundo real» de poner en funcionamiento una plataforma «instantánea, pública y mundial»,⁴⁹ escribió Dorsey ese marzo. Admitía que aquello había provocado daños reales. En entrevistas, empezó a plantear por voluntad propia ideas heréticas que otros directores ejecutivos de empresas tecnológicas seguían rechazando con fervor: tratar de maximizar la participación es peligroso; los «me gusta» y los retuits fomentan la polarización. La compañía, dijo, modificaría sus sistemas para promover conversaciones «sanas» en lugar de conversaciones que enganchasen a los usuarios. Contrató a destacados expertos y grupos de investigación para desarrollar nuevas características o elementos de diseño para lograrlo.

Pero casi ninguna de esas iniciativas llegó a despegar. Muchos de los expertos externos, después de meses de retrasos o cambios repentinos de las políticas, abandonaron frustrados.⁵⁰ No estaba claro si el experimento de Dorsey para reinventar Twitter había fracasado porque su atención se fue a otra cosa; porque los inversores, cada vez más hostiles, presionaron Twitter para que se centrara en el crecimiento o porque las soluciones resultaron ser inaceptables para una compañía atrapada aún en la mentalidad de Silicon Valley. Informaciones de trabajadores de Twitter hacen pensar que probablemente fuera una combinación de los tres factores.

En YouTube, mientras tanto, todo iba como siempre. Los sistemas de la plataforma seguían creando comunidades marginales con alta participación. Ese año, la psicóloga cognitiva Asheley Landrum descubrió una en Denver, en un congreso para personas que creían que la Tierra era plana.⁵¹ El terraplanismo, que llevaba siglos extinguido, de pronto estaba resurgiendo. Profesores de ciencia contaban que había estudiantes que cuestionaban sus explicaciones

con cálculos de curvatura y diagramas corregidos. Una estrella de la NBA se declaró públicamente terraplanista. Y lo mismo hizo un rapero. Landrum se paseó por el congreso de Denver con la esperanza de entender por qué eso estaba sucediendo. Los resultados no fueron ambiguos. De los treinta asistentes entrevistados, veintinueve dijeron que habían conocido el terraplanismo y se habían convencido de esa idea en YouTube. Al trigésimo entrevistado lo había convencido su hija, que había encontrado el movimiento en YouTube.

Guillaume Chaslot, el extrabajador de YouTube, se había percatado de que esos vídeos tomaban protagonismo en la plataforma ya cuando trabajaba para la compañía, y había alertado a nivel interno de que aquello apuntaba a algo peligroso. Como un colorante médico utilizado para hacer visible una infección, su presencia suele indicar que una red social fomenta creencias extremistas de toda clase, algunas de ellas peligrosas. «Me dijeron: “La gente clicla vídeos terraplanistas; por tanto, quieren vídeos terraplanistas” —recordaba Chaslot—. Y mi argumento era: no, no es que, porque alguien clicó un vídeo terraplanista, eso significa que quiere que le mientan. El usuario simplemente tiene curiosidad, y hay un título que es un ciberanzuelo. Pero para el algoritmo, cuando ves un vídeo, eso quiere decir que apoyas lo que en él se dice.»

Durante mucho tiempo habían aparecido vídeos terraplanistas con miniaturas y títulos provocadores en las recomendaciones de la plataforma: «Doscientas pruebas de que la Tierra no es una bola que da vueltas», «Pistas de que la Tierra es plana», «¡La Estación Espacial Internacional no existe!». La gente clicaba por curiosidad. Al algoritmo no le importaba por qué los usuarios veían los vídeos, solo que lo hacían. Y al hacerlo, recomendaba esos vídeos cada vez más.

YouTube, al mostrar a los usuarios muchos vídeos seguidos difundiendo todas las mismas ideas, incide con especial intensidad sobre dos de nuestros puntos débiles cognitivos: que la exposición reiterada a una afirmación,⁵² así como la impresión de que la afirmación es aceptada de forma generalizada,⁵³ hacen que parezca más cierta de lo que de otro modo pensaríamos. La mayoría de los usuarios, claro está, probablemente rechazan los vídeos de conspiraciones. Pero a una escala de miles de millones, esos métodos atraviesan suficientes defensas entre las personas susceptibles para terminar captando a miles de conversos para incluso la más absurda

de las causas. O la más peligrosa.

UN GRAN DESPERTAR

En octubre de 2017, dos meses después de la concentración «Unir a la derecha» en Charlottesville, apareció una breve publicación en el tablón de política de 4chan escrito por el usuario Q Clearance Patriot. El usuario daba a entender que era un oficial de la inteligencia militar que participaba en una operación para detener a los miembros del Pizzagate, la conspiración según la cual dirigentes demócratas operaban una red internacional de tráfico de menores. El usuario, que afirmaba poseer una autorización del Gobierno de nivel Q, anunció que la extradición de Hillary Clinton ya estaba «en marcha» y que se había movilizado a la Guardia Nacional para contrarrestar los «disturbios masivos que se organizarían en contra de esa acción». Pero fue su segunda publicación la que marcó el estilo por el que se conocería a Q:

Mockingbird

HRC detained, not arrested (yet).

Where is Huma? Follow Huma.

This has nothing to do w/ Russia (yet).

[Mockingbird

HRC retenida, no detenida (todavía)

¿Dónde está Huma? Seguid a Huma.

Esto no tiene nada que ver con Rusia (todavía).]⁵⁴

La publicación seguía a lo largo de otras veinte líneas. Con referencias lo bastante crípticas para que los usuarios tuvieran la sensación de que estaban descifrando un código secreto y a la vez lo suficientemente obvias para asegurarse de que lograrían entenderlas. Como la primera página de una novela de espías, sentaba las bases de una trama que mantendría su regularidad durante miles de «gotas de Q», como los fans llamaban a las publicaciones. Trump y sus generales se estaban preparando para detener a miles de conspiradores

demócratas e imponer un Gobierno militar en lo que sería una jornada sangrienta y gloriosa de ajuste de cuentas.

A lo largo de los siguientes meses, la historia fue subiendo de tono. Se detendría a decenas o cientos de miles de personas: élites culturales, financieros, burócratas del «Estado profundo» y espías. Durante generaciones, esa camarilla había controlado en secreto la vida estadounidense, responsable de toda clase de atrocidades, desde el Pizzagate hasta la injusticia del orden económico. Ahora serían ejecutados, en un National Mall convertido en un campo de exterminio. Algunos ya habían sido sustituidos por dobles, un primer golpe discreto del plan de Trump.

Los seguidores tenían algo más que una historia. El QAnon, como se llamaba a sí mismo el movimiento, se convirtió en una serie de comunidades virtuales donde los creyentes se congregaban para analizar las publicaciones de Q. Buscaban pistas y significados ocultos en las gotas y, a petición de Q, en detalles de la política diaria. ¿Era el comentario informal de Trump sobre el Congreso una referencia en clave de la inminente purga? ¿La extraña colocación de su mano, una señal para el batallón 101 de las Fuerzas Aerotransportadas? Era un juego interminable, una actividad de grupo para estrechar los vínculos sociales que tenía repercusiones en la vida cotidiana.

Desde hace mucho tiempo, los grupos extremistas reclutan miembros con la promesa de satisfacer su necesidad de encontrar una razón de ser y un sentimiento de pertenencia. A veces el odio no es más que el pegamento que une. «La camaradería social —ha escrito la socióloga Kathleen Blee— puede coexistir con el odio (o incluso sustituirlo) como el motivo para participar en actividades racistas organizadas.»⁵⁵ Las conspiraciones, por su parte, prometen el fin de los sentimientos de impotencia en medio de un mundo caótico e incomprensible.⁵⁶ Quizás las fuerzas del mercado te han quitado el trabajo. Una enfermedad o una catástrofe repentina han puesto tu vida patas arriba. Los cambios sociales han hundido tu sentido del orden legítimo de la sociedad. (Muchos miembros de QAnon eran adinerados, pero habían caído en un pánico desorientador por culpa de, por ejemplo, la elección de un presidente negro o el aumento de la diversidad.)⁵⁷ Las conspiraciones insisten en que los hechos, en lugar de ser incontrolables o impersonales, forman parte de un complot oculto cuyos secretos pueden ser descifrados. Redefinir el caos como

un orden,⁵⁸ contando a los creyentes que solo ellos poseen la verdad, restablece su sentimiento de autonomía y control. Por eso los adeptos al QAnon suelen repetirse unos a otros su mantra reconfortante: «Confía en el plan».

Los expertos sobre el extremismo especularon durante largo tiempo que muchas de las publicaciones de Q —4.000 en total, desveladas a lo largo de tres años— eran en realidad obra de Ron Watkins, un programador de treinta años que poco antes había asumido la gestión de 8chan, el foro escindido de 4chan. Watkins incluso pareció insinuar que sí en un documental de 2021, donde decía al entrevistador: «Fueron básicamente tres años de formación en inteligencia, enseñando a gente normal cómo hacer tareas de inteligencia. —Aunque añadía—: Pero nunca he sido Q». Sin embargo, la mayoría de sus seguidores desconocían las pruebas de la identidad real de Q o quizás, de alguna forma, no querían saberlo. El atractivo estaba en lo que les ofrecía la historia, no en su autoría o en su verdad objetiva. Muchos investigadores que rastrearon QAnon consideraban incluso que la identidad de Q no era más que un detalle menor. A pesar de la manipulación maquinadora de Q, eran los usuarios, las plataformas y las tendencias de enlace de ambos lo que propulsaba el movimiento.

QAnon, como buena parte de la tecnocultura que lo precedió, se propagó enseguida de los márgenes a las plataformas mayoritarias. Los sistemas de Facebook y YouTube introdujeron QAnon en la estela de las recomendaciones extremistas y conspirativas.⁵⁹ Al principio, era solo un nodo más. Pero sus afirmaciones, cada vez más subidas de tono, y su historia totalizadora le permitieron absorber otras conspiraciones, lo cual lo convirtió en un punto de fuga —como había sido Alex Jones tiempo atrás— de comunidades que no estaban vinculadas entre ellas, desde los antivacunas hasta los que se preparan para la guerra racial o los paranoicos antigobierno. Su política vehementemente pro-Trump y sus claros llamamientos a la violencia lo convirtieron en un socio natural, y finalmente en un elemento habitual, de la derecha digital. Pero, por encima de todo, tenía una capacidad infinita para captar la atención de los usuarios, el rasgo principal que las plataformas quieren maximizar.

Cuando los estadounidenses se dieron cuenta de que eso era algo peligroso, los grupos de Facebook sobre QAnon ya tenían millones de

miembros, los vídeos de QAnon recibían millones de visitas en YouTube y las cuentas sobre QAnon en Twitter organizaban campañas de acoso masivas contra famosos a quienes acusaban de estrambóticas tramas de canibalismo. Una aplicación que agregaba gotas de Q se convirtió en una de las descargas más populares en la App Store.⁶⁰ Un libro, *QAnon: An Invitation to a Great Awakening* [QAnon: Invitación a un gran despertar], escrito por un colectivo de adeptos anónimos, llegó al número dos de la lista de los más vendidos en Amazon.⁶¹ Los miembros se pasaban todo el tiempo inmersos en la comunidad, dedicando varias horas al día a los chats de vídeo y a los hilos de comentarios que se habían vuelto su mundo.

En mayo de 2018, un *youtuber* vinculado a QAnon entró de forma violenta en una fábrica de cemento de Arizona que, según afirmaba él, era el centro de una red de tráfico de menores y emitió el enfrentamiento en directo para cientos de miles de espectadores en Facebook. Al mes siguiente, otro adepto, con un rifle AR-15 en las manos, bloqueó los dos carriles del tráfico de la presa de Hoover con un camión blindado de fabricación casera exigiendo que el Gobierno publicara un informe que Q había mencionado en el mensaje del día anterior. Otro fue detenido en Illinois con materiales explosivos que pretendía detonar en el edificio del capitolio del estado, con el objetivo de atacar a satanistas abusadores de niños que creía que habían tomado el poder. Y otro condujo hasta el domicilio de un capo de la mafia de Staten Island, al que creía que Q había señalado como un conspirador del Estado profundo, y lo mató a tiros. El FBI, en un comunicado interno, identificó QAnon como una potencial amenaza de terrorismo nacional. Aun así, Facebook, YouTube y Twitter, como habían hecho con Alex Jones, en buena medida se negaron a tomar medidas y siguieron permitiendo que el movimiento creciera, potenciado por los sesgos de sus plataformas.

Casi toda la extrema derecha digital parecía atraída por el movimiento. Así como muchos agentes del orden, atraídos por sus promesas de establecimiento de un orden firme y represalias contra los liberales.⁶² El jefe del sindicato del departamento de policía de Nueva York daba entrevistas en televisión con su taza de QAnon en el plano. El vicepresidente Mike Pence fue fotografiado junto a un agente de las fuerzas especiales de Florida que llevaba una insignia con una Q. El movimiento migró a Instagram, que promovió QAnon con tal

intensidad que también cayeron muchas de las madres aficionadas al yoga y los *influencers* de estilo de vida que dominan la plataforma.⁶³

Pero, a pesar de la sensación de autonomía, seguridad y comunidad que ofrecía QAnon, eso tenía un coste: un aislamiento demoledor. «Voy a ser sincero, patriotas —escribió un popular seguidor de QAnon en un tuit muy compartido dentro del movimiento—. Cuando empecé a meterme en Q y los patriotas y el movimiento, me sentí muy solo porque la mayoría de la gente que conocía, aunque fueran conservadores, creen que todo esto es una teoría de la conspiración. ¡Este es el único sitio en el que me siento acogido y como en casa, aquí con todos vosotros!»⁶⁴ Muchas personas respondieron con historias parecidas. «Estoy en el mismo barco. Literalmente, sin amigos ni familia con quienes hablar. Gracias a Dios, mi esposa y yo seguimos al conejo juntos, porque eso es lo único que tenemos. Mi familia política me abandonó como si yo no les importara en absoluto tras la muerte de mi padre. Así que estoy agradecido por todos los que estamos en este barco.»⁶⁵ Esa era una de las cosas que hacían que QAnon radicalizara tanto. Unirse al movimiento a menudo agravaba la propia sensación de aislamiento y de ir a la deriva que había llevado a las personas hasta él. Sin ningún sitio al que acudir y ahora doblemente necesitados de consuelo, los seguidores se entregaban a la causa con más devoción todavía.

En 2019, uno de los investigadores de Facebook creó una serie de cuentas de prueba que simulaban la experiencia media de ciertos usuarios arquetípicos.⁶⁶ Uno de ellos era «Carol Smith», una madre ficticia de Carolina del Norte. El investigador introdujo a Carol en páginas sobre paternidad, cristianismo y política conservadora, y luego esperó para ver adónde la llevaba Facebook. Dos días después, la plataforma la arrastró hacia QAnon. Pasados cinco más, la guio, según escribió el investigador en un informe interno, por «una mezcla de contenido extremo, conspiratorio y gráfico».

En otro informe, también interno, se constataba que la mitad de los usuarios de Facebook vinculados con QAnon, que ascendían por lo menos a los 2,2 millones, se habían unido al movimiento, como en el caso de YouTube, mediante «grupos de entrada».⁶⁷ Eso confirmaba las advertencias de DiResta —que venía haciendo desde que había descubierto a los antivacunas años antes— de que esa característica de la plataforma era un vehículo para la radicalización. «No puedo dejar

de insistir en lo catastróficos que son los Grupos —tuiteó en 2018, cuando aumentaban las pruebas—. El motor de recomendación de los Grupos es una matriz de correlación de conspiraciones. Arrastra a los usuarios propensos al contenido extremista o polarizador hacia grupos cerrados y luego secretos. FB no tiene ni idea de lo que ha creado con esta herramienta.»⁶⁸

Sin embargo, Facebook no hizo más que ahondar su compromiso con la herramienta que sus propias investigaciones habían demostrado que era responsable de la mayor parte del crecimiento de QAnon. Dijo que cien millones de usuarios, que esperaba que aumentasen a los mil millones, estaban activos en grupos que habían surgido donde pasaban la mayoría de su tiempo. No obstante, incluso algunos trabajadores de Facebook hacían sonar las alarmas. Otro informe interno, filtrado después al *Wall Street Journal*, advertía: «El 70 % del top cien de los Grupos Cívicos de Estados Unidos más activos se consideran no recomendables por aspectos como el odio, la desinformación, la intimidación y el acoso».⁶⁹ Un ingeniero que había trabajado en la herramienta alertó a un periodista de *The Verge* de que allí era «donde empieza la generación de la burbuja». La práctica de empujar a los usuarios hacia grupos que se seleccionaban para captar su atención, añadió el ingeniero, se había vuelto «muy peligrosas».⁷⁰

NIHILISMO DIGITAL

En su primer año de universidad, Adam, el usuario de 4chan de toda la vida de Dallas, entregó su lealtad a 8chan. Era el año 2014, QAnon y muchas más cosas estaban por venir, pero la promesa de 8chan de acoger a los seguidores del Gamergate, expulsados incluso de 4chan, había distinguido la plataforma como el último hogar auténtico de la libertad de expresión. Sus debates sin restricciones y su contenido extremo le habían dado la reputación de ser la taberna pirata de la web social, repleta de contenido imposible de encontrar en otras partes. «Veía mucha mierda —me contó Adam—. Cáteles decapitando personas, porno horrible que no debería existir; no quiero ni entrar en detalles. Pero lo ves cien veces y entonces los problemas de la vida real empiezan a parecer menos impactantes.»

8chan era ir más allá de los límites como en 4chan pero sin ni

siquiera las reglas de 4chan. Los miembros de 8chan llegaban a nuevos extremos como una defensa colectiva ante la anomia que, en tanto que marginados de los marginados, inadaptados dentro de los inadaptados, ponían en el centro de su identidad compartida. «Su teoría sobre lo que hacían allí, sobre lo que sacaban de allí, era que estaban aprendiendo a no reaccionar ante las personas que pulsaban sus botones emocionales o ideológicos —escribió sobre sus miembros Dominic Fox, ingeniero del *software*—. El mundo real era un lugar duro e insensible, y cualquier persona que fingía que le importaba (o que necesitaba que le importase) estaba, por definición, atrapado en un engaño, en una suerte de estafa.» Así pues, según Fox, los miembros de 8chan pensaban que «la única forma de liberarse de ese control era mirar memes racistas, fotos de accidentes de tráfico, pornografía horripilante y así sucesivamente hasta que uno pudiera sobrellevarlo con total serenidad». Y consideraba que era una cultura de «desensibilización deliberada».⁷¹

Con el tiempo, así como la transgresión de 4chan se convirtió en el elemento distintivo del endogrupo, eso mismo sucedió con la desensibilización en 8chan. Tolerar cosas demasiados chocantes o insoportables para los externos a la plataforma era una forma de demostrar que pertenecías al grupo. Tú también te habías vuelto insensible, tenías los ojos abiertos, eras un soldado de una hermandad de vacuidad. En las secciones más duras, los usuarios que encontraban confort y comunidad en esa práctica buscaban tabúes cada vez más fuertes hasta que llegaban, quizás de forma inevitable, al más extremo de todos: el asesinato masivo.

Los usuarios inventaban memes y chistes internos elaborados en torno a la alabanza (y la compartición de imágenes gráficas) del Holocausto, de matanzas masivas y, sobre todo, por su proximidad en el tiempo, del genocidio de musulmanes bosnios en 1995. El atractivo de celebrar un genocidio o de elogiar a un tiroteador estaba, en cierto modo, en que los demás no lo harían. Eso te acreditaba como un auténtico miembro de 8chan: alguien a quien le gusta la transgresión y que se ha tomado la pastilla roja, que está comprometido con los otros miembros sin importarle el coste que eso tenga para la reputación y que hace caso omiso de las triviales exigencias de un mundo físico que, en realidad, le aterroriza. La cultura predominante no era de ultraderecha de por sí, aunque el lenguaje del nazismo y del

genocidio racial estaba por todas partes. La culminación de muchos años cruzando la línea y haciendo provocaciones, expresando un odio irónico que se había vuelto sincero, era más bien una suerte de nihilismo violento. Los usuarios se alentaban unos a otros, medio en broma medio en serio, a castigar al mundo, a hacerles daño a Ellos en nombre de Nosotros. A matar de forma indiscriminada.

Era fácil no darle importancia, a menos que uno hubiera estado prestando atención. En las redes sociales, había comunidades de anomia y crisis —íncels de Reddit, grupos de Facebook sobre Q, la extrema derecha de YouTube— que habían ido enconándose hasta el extremo de pasar a la acción. Aunque solo una pequeñísima minoría llevaba a la práctica las amenazas, así era como había funcionado siempre el extremismo violento. No había motivos para pensar que en el caso de 8chan no iba a suceder lo mismo.

En marzo de 2019, mientras empezaba el rezo de la tarde en una mezquita de planta baja en un frondoso barrio de Christchurch, en Nueva Zelanda, un hombre desconocido entró en el recinto en coche. Llevaba una escopeta y tenía un móvil enganchado en el pecho. Se acercó a cuatro hombres que estaban charlando en la entrada. Uno lo saludó: «Hola, hermano». Él agarró la escopeta y mató a los hombres con varios disparos.

Luego entró por los pasillos estrechos y de techo bajo de la mezquita, donde se habían congregado casi doscientos fieles, y empezó a disparar enseguida. Fue baleando a un feligrés tras otro mientras los acorralaba hacia la pared trasera. Entonces tiró la escopeta, se le había acabado la munición. Agarró un fusil AR-15, con una luz estroboscópica que centelleaba desde la boca del arma, y disparó metódicamente mientras las familias permanecían paralizadas por el miedo o intentaban huir. Naeem Rashid, un maestro de cincuenta años y padre de tres hijos, se abalanzó contra el desconocido, que lo mató de un tiro. El hombre armado no dijo ni una palabra. En lugar de eso, por un altavoz portátil que llevaba atado al pecho se oían a todo volumen himnos nacionalistas. Fue paseándose por la mezquita y disparando a todas las personas que encontraba, luego regresó a la sala de oración principal, donde disparó de nuevo contra los cuerpos de víctimas ya muertas o moribundas. Cuando salió del edificio, solo cinco minutos después de su llegada, había asesinado a cuarenta y tres personas. La más mayor tenía setenta y siete años; la

más joven, tan solo tres. Mientras se dirigía a su coche, se detuvo junto a una mujer tumbada en el suelo que sangraba mucho pero estaba viva. El hombre se inclinó sobre ella. En ese momento, doscientas personas lo estaban viendo en directo por Facebook, siguiendo la retransmisión hecha con el móvil que llevaba en el pecho. El hombre escuchó cómo la mujer le suplicaba que la dejase vivir y luego la mató.

Condujo —girando con brusquedad entre carriles a casi ciento cuarenta y cinco kilómetros por hora y disparando desde la ventanilla de su coche— hasta otra mezquita, donde asesinó a siete personas más. Abdul Aziz Wahabzada, un inmigrante de cuarenta y ocho años de Afganistán, interrumpió el ataque arrojándole un datáfono y luego agachándose detrás de un coche. Mientras el hombre se rearmaba, Wahabzada salió corriendo de su escondite, recogió una escopeta sin munición que el pistolero había tirado y le apuntó. Nervioso, el tiroteador se marchó en su coche mientras Wahabzada lanzaba la escopeta contra la ventana trasera del vehículo, que se rompió. Mientras el pistolero iba de camino a una tercera mezquita, una pareja de policías rurales, que estaban de visita en la ciudad para una sesión de formación, embistió el coche del hombre, que chocó contra el bordillo, y lo detuvieron.

Al cabo de unas horas, se constató que el asesino, un australiano de veintiocho años llamado Brendon Tarrant, representaba una nueva clase de extremismo violento; la oscuridad más absoluta de la web social profunda se había manifestado. «Pues nada, chavales, ha llegado la hora de dejar de hacer publicaciones provocativas y de hacer una publicación sobre una iniciativa en la vida real», había escrito en el foro político de 8chan horas antes del ataque. Puso un enlace a su Facebook Live, animando a los usuarios a seguir su «ataque contra los invasores». Y adjuntó un archivo de texto de setenta páginas explicándose, un texto que también publicó en Twitter. «Sois todos unos tíos geniales y los mejores amigos que un hombre podría pedir —escribió—. Por favor, colaborad difundiendo mi mensaje, haciendo memes y colgando publicaciones provocativas como soléis hacer.»

Los usuarios que lo siguieron en directo, los 4.000 que vieron su vídeo antes de que Facebook lo eliminara y otros muchos que vieron los millones de subidas posteriores que circularon por internet, al

parecer le oyeron decir, momentos antes de empezar a matar.⁷²«Recordad, chavales, suscribíos a PewDiePie». Era, en una pista de los retorcidos motivos de Tarrant, una broma interna de las redes sociales. Los fans de PewDiePie —el *gamer* más popular de YouTube, con más de cien millones de suscriptores en su canal— habían llenado internet poco antes —y, para gastar una broma, habían atascado faxes de oficinas— con esa misma frase.⁷³El *youtuber* también había sido objeto de acusaciones (en buena medida desatinadas) de flirtear con el nacionalismo blanco. La referencia de Tarrant pretendía ser un troleo a los medios de comunicación, pues los retaba a vincular su violencia con PewDiePie, y un guiño a sus compañeros troles de internet que entenderían la broma.

El terrorismo es violencia destinada a amenazar a una comunidad más amplia con finalidades políticas o simple maldad. Pero suele ser también un espectáculo para el endogrupo y un acto de solidaridad con él. Por eso los militantes de Estado Islámico, sobre todo los «lobos solitarios» radicalizados en internet que nunca habían conocido en persona a otro miembro, inundaban los foros yihadistas de manifiestos y mensajes en vídeo presentándose como mártires que solo tendrían sentido para otros creyentes.

Por este mismo motivo las últimas palabras escritas de Tarrant expresaron sobre todo afecto por la comunidad por la que, según aclaraba, hacía lo que estaba haciendo. En su documento, Tarrant escribió que había descubierto el «etnonacionalismo» en *Spyro: Year of the Dragon*, un videojuego infantil. Daba las gracias a Candace Owens, una popular *youtuber* de la derecha alternativa, por enseñarle a emplear la violencia. En un párrafo decía presumiendo: «Tengo formación en la guerra de gorilas y soy el mejor francotirador de todas las fuerzas armadas de Estados Unidos». Todo eso eran chistes internos de la web social.

Pero por cada chiste o trol había páginas de invocaciones aparentemente sinceras a conspiraciones de la extrema derecha, eslóganes nazis y llamamientos a una guerra racial mundial para expulsar y exterminar a los no cristianos y no blancos. Por encima de todo, Tarrant envolvía su violencia con bromas internas que para su comunidad se habían convertido en algo terriblemente sincero, materializaciones de su deriva desde el extremismo irónico hasta un extremismo que estaba solo disfrazado de ironía.

Durante su retransmisión en directo, mientras conducía, puso una canción de un vídeo pixelado de 1992 de un ultranacionalista serbio en que se elogia a Radovan Karadžić, un criminal de guerra responsable del genocidio bosnio. El vídeo había sido durante mucho tiempo un meme de 4chan, llamado «remove kebab» ('eliminar kebab'), que, para algunos, pasó de broma a señal auténtica de apoyo al genocidio de musulmanes. Tarrant también había escrito «remove kebab» en el lateral de uno de sus fusiles.

En 8chan, los usuarios, al verlo en directo, estaban eufóricos.⁷⁴

¡JAJAJA PUSO REMOVE KEBAB DURANTE EL TRAYECTO! ¡ME PARTO DE RISA!

Madre mía. El OP armó una buena acabo de verle matando a un montón de putos moros. [OP significa *original poster* ('publicador original'), en referencia a Tarrant.]

!!!HOSTIA PUTA!!! !!!LOS DÍGITOS DE DIOS!

Algunos se animaban unos a otros a seguir su ejemplo y, como expresó uno, a «redimir a su nación».

Transcurridos varios meses, Fredrick Brennan, el fundador de 8chan y su administrador hasta 2016, dijo que el sitio debía cerrarse.⁷⁵«No está aportando nada positivo al mundo —dijo a Kevin Roose, columnista sobre tecnología del *New York Times*—. Es algo completamente negativo para todo el mundo salvo para sus usuarios. Bueno, en realidad no. También es negativo para ellos. Solo que no se dan cuenta.»

En pocos meses, se anunciaron en el foro otros dos asesinatos masivos con motivaciones supremacistas blancas. Un usuario de diecinueve años, después de publicar sus intenciones, entró con un AR-15 y cincuenta cartuchos de munición en una sinagoga de California y disparó contra cuatro personas, mató a una; acto seguido se le atascó el fusil y huyó. Luego un chico de veintiún años en El Paso mató a veintitrés personas en un supermercado Walmart, la mayoría latinos. La policía dijo que, antes del ataque, había publicado un largo mensaje en 8chan relatando conspiraciones de extrema derecha y guerra racial que creía que justificaban su violencia.

A pesar de los vínculos con 8chan, cuando los investigadores del Gobierno de Nueva Zelanda terminaron su análisis, que duró años, de cómo se había producido la masacre de Christchurch, la mayor culpabilidad recaía, según indicaron, en YouTube.⁷⁶ Tarrant había aprendido en esa plataforma cómo modificar las armas que utilizó en el ataque. Había donado dinero a *youtubers* de la derecha alternativa. En los días previos al ataque, había dejado la plataforma en modo «reproducción automática», lo cual había permitido que el algoritmo lo condujera de un vídeo al siguiente. En los demás sitios por los que se movía de internet, ya fuera en grupos de Facebook o en foros de chats de amantes de los videojuegos, solía publicar enlaces a vídeos de YouTube de extrema derecha.

Los investigadores, citando entrevistas y reconstrucciones forenses del historial web de Tarrant, llegaron a la conclusión de que «YouTube fue, en su caso, una fuente de información e inspiración mucho más relevante»⁷⁷ que cualquier otra plataforma. Según el informe, YouTube fue tanto el hogar digital como el motor de su radicalización. Cuando los investigadores preguntaron a personas de la zona sobre el ataque, ellas también mencionaron YouTube: su rol en la radicalización de la extrema derecha, su promoción de contenido extremista. Algunos dijeron que el algoritmo de YouTube incluso les había recomendado el vídeo, supuestamente prohibido, de la matanza de Tarrant. Jacinda Ardern, primera ministra de Nueva Zelanda, declaró que, entre los cientos de páginas de conclusiones del informe, el papel de YouTube a la hora de fomentar el odio y las conspiraciones «tenía una especial relevancia». Y añadió: «Este es un punto que tengo previsto tratar con los directivos de YouTube».⁷⁸

A decenas de miles de kilómetros de allí, en una oscura habitación de las afueras de Dallas, Adam, sin tener intención de hacerlo, había visto la matanza en 8chan poco después de que tuviera lugar. «Era espantoso —recordó—. A mí, que me considero una persona desensibilizada, aquel vídeo me horrorizó.»

Eso se produjo en un momento de debilidad para él. Llevaba años en comunidades íncel, siguiéndolas en YouTube y en 8chan mientras descendían —las comunidades y el propio Adam— hacia la rabia y la desconfianza. Pero había conocido a una chica en un grupo de Facebook, una página de compartición de memes que el algoritmo de Facebook le había recomendado. Adam había publicado un hilo en el

que se ofrecía a dibujar a quien le mandase una foto suya. La chica había aceptado, pero, cuando intercambiaron información de sus perfiles, había utilizado el acceso para dibujarlo ella a él. La alondra bajó la guardia. Todo lo que había aprendido en internet le decía que las mujeres solo lo denigrarían y le harían daño. Pero ahí había una prueba de que alguien lo había visto, se había interesado por él por voluntad propia y no lo veía como alguien repugnante sino un poco tontorrón y amable. Tras meses hablando por internet, ella hizo un trayecto de tres horas en coche hasta Dallas. Al principio, se mostraban tímidos. «Eso sería unos cuatro o cinco días antes de que se derramara la tetera», dijo el chico. Pero después mantuvieron el contacto. Ella empezó a sacarlo de la madriguera. Luego ocurrió lo de Christchurch y él se hartó.

Poco después, visitó la página de Twitter de Brianna Wu, la desarrolladora de videojuegos y activista que, durante el Gamergate, había sufrido años de acoso y amenazas. «hola, quizás te parecerá una tontería —le escribió en un mensaje privado—, pero bueno, solo quería decirte que hace algunos años yo era una persona bastante nerviosa que se vio atrapada en medio del odio de 4chan contra ti». No mencionaba qué había hecho en concreto. Pero decía que se había visto absorbido por una «enorme mentalidad colectiva» de odio antifeminista, contra ella y otras personas, en 4chan y en Facebook. «las cosas han cambiado y te apoyo a ti y tu campaña», escribía, en referencia a la candidatura de Wu para el Congreso. «Pensaba que merecía la pena contártelo, más que nada para mostrarte mi apoyo y que esas personas pueden cambiar a mejor.»

Wu llevaba tiempo recibiendo cada vez más mensajes de ese tipo: de exparticipantes en el Gamergate que, tras años metidos en el embudo que los adentraba más y más en el extremismo, habían pedido ayuda. «En realidad me parece que son personas que buscan su lugar en el mundo —me contó ella—. Entiendo que seas infeliz y canalices el odio que llevas dentro hacia fuera, hacia otras personas, y que contraataques, sobre todo cuando eres joven. Así que trato de recordar cuando tenía veintipocos años. Estaba hecha un lío.»

Contestó a Adam: «Eso significa mucho para mí, ¡gracias!». Tras algunos mensajes, le pidió permiso para publicar una captura de su conversación en Twitter quitando su nombre. Él aceptó, tenía la esperanza de que pudiera ayudar «a más personas con las que crecí

hablando por internet» a cambiar. El tuit de Wu recibió quizás una docena de respuestas, la mayoría positivas. Fue probablemente la cuarta o quinta publicación más comentada de ese día. Pero para Adam la sensación de afirmación fue desbordante.

«ver todas las respuestas al pantallazo que compartiste me emocionó mucho —le dijo por privado—. creo que nunca había visto tantas personas tan contentas por algo que yo hubiese dicho.»

Los nuevos líderes supremos

UN GOBIERNO ESCONDIDO

Se suponía que el trabajo de Jacob iba a ser fácil. El moderador de Facebook, quien tiempo después leería mi artículo sobre Sri Lanka y me mandaría montones de documentos internos para alertar al mundo ante lo que él veía como una peligrosa negligencia empresarial, se sentaba delante del ordenador, donde se le mostraban contenidos de Facebook, Instagram o WhatsApp. Revisaba las directrices para ver si la publicación era permisible. Clicaba para autorizarla, eliminarla o consultarla con un superior. Y siguiente.

Al principio, esas directrices eran una lista sencilla, en la cual se prohibían los insultos racistas, la desnudez y poco más.¹ Pero en 2018 las normas parecían las instrucciones para operar un reactor nuclear, si cada página hubiese sido escrita por un autor distinto, todos ellos desconocedores de cómo su página encajaba en el conjunto. Se extendían durante más de 1.400 páginas (probablemente muchas más, si se tienen en cuenta archivos específicos de una región a los que el equipo de Jacob no tenía acceso). Y, aun así, se pretendía que Jacob y sus compañeros de trabajo, la mayoría de ellos exoperadores de servicios de atención al cliente, conocieran y aplicaran esas normas para tomar cientos de decisiones de gran calado día tras día. Su oficina, con una plantilla de algunas decenas de personas que revisaban contenido en varias lenguas y de varias regiones, era una de las muchas que hay esparcidas por el planeta. Un enorme archipiélago de miles de moderadores distribuidos en montones de oficinas, que compartían poca comunicación o coordinación más allá de lo que venía de la remota sede central de Facebook.² Árbitros ocultos que separaban lo prohibido de lo permitido, que sin ser vistos determinaban la política y las relaciones sociales en todos los países de la Tierra.

Así fue como Jacob, cansado de sentirse cómplice del veneno que

veía cómo iba surgiendo en las plataformas propiedad de Facebook — una ola de mentiras sectarias que alimentaba la violencia comunitaria en Irak, una conspiración de provocaciones raciales que se formaba en Israel o India— y de no encontrar respuesta por parte de los jefes de la agencia subcontratada a quienes había intentado alertar, terminó por contactar conmigo. Las normas que regulan las redes sociales, escribía en su primer mensaje, «no están adaptadas a la realidad del contenido que vemos». Y también le preocupaba otra cosa. Facebook, decía, había descargado esa «labor sensible, que tiene consecuencias para la seguridad internacional», a multinacionales como aquella para la que él trabajaba, una «empresa con ánimo de lucro a la que solo le importa maximizar la productividad». Por muy malas que fueran las normas, el modelo de negocio de reducir los gastos al mínimo, maximizar los beneficios y actuar con secretismo solo agravaba mucho más todavía los errores en la moderación de contenido.

A lo largo de los siguientes meses, Jacob y yo hablamos con regularidad por aplicaciones seguras, y al final tomé un avión para ir a recoger los manuales que él había sacado a escondidas de los sistemas informáticos de su empresa. Los archivos prometían aportar algo poderoso: pruebas de que Facebook se había integrado, aunque fuera a regañadientes, en la gobernanza de un mundo cuya política y relaciones sociales pasaban cada vez más por sus sistemas. Y eso con enorme detalle sobre cómo lo habían hecho.

Mientras estábamos hundidos en un sofá viejo en la casa de dos habitaciones de Jacob, hecha de bloques de hormigón, bebiendo de una botella de dos litros de refresco que había comprado para la ocasión, me contó las espirales de odio e incitación que su equipo había sido forzado a dejar en la red por unas normas que ellos habían calificado una y otra vez como insuficientes y defectuosas. «Al fin y al cabo —dijo—, te ves forzado a seguir las normas de la compañía si no quieres perder el trabajo.» Pero tenía remordimientos por aquellas decisiones. «Tienes la sensación de que has matado a alguien por no haber actuado.»

Él sabía que no era nadie, una mera pieza más de los miles que constituyen el engranaje mundial de Facebook. Y, a pesar de sus preocupaciones, aún tenía fe en la compañía que tanto había prometido al mundo; desde luego, eso no era nada más que un pequeño fallo en el camino para hacer realidad el gran proyecto de

Mark Zuckerberg. Incluso cuando yo estaba en su casa, leyendo con detenimiento documentos cuya publicación él sabía que avergonzaría a Facebook, Jacob se veía como su aliado. Desvelando los fallos burocráticos que temía que estaban siendo un lastre para la compañía, la ayudaría a llevar a cabo la revolución tecnológica en la que había depositado sus esperanzas.

El mamotreto no parecía un manual para regular la política internacional. No había un archivo maestro o una guía general, solo decenas de presentaciones de PowerPoint y hojas de cálculo de Excel desvinculadas entre ellas. Sus títulos, aunque eran burocráticos, daban pistas del ámbito de actuación de Facebook. Violencia plausible: estándares de implementación. Productos regulados. Sri Lanka: discursos de odio. Terrorismo: política de perfiles falsos/no reales. Organizaciones de odio de los Balcanes occidentales y sus cifras.

Las normas abarcaban casi cualquier materia que la compañía podía querer que sus moderadores tuviesen en cuenta. Y tenían la intención de dar, para cualquier eventualidad en cualquier país o región, unas instrucciones precisas y maquinales para decidir qué eliminar y qué autorizar. Era un proyecto de una ambición imponente. El objetivo era reducir las preguntas con mucho contexto que incluso un equipo de abogados especializados tendría dificultades para analizar —qué constituye una amenaza, cuándo una idea transmite odio, cuándo un rumor es peligroso— a una cuestión de blanco o negro tan meridiana que cualquier moderador pudiera tomar decisiones sin tener que hacer una reflexión propia. Muchos de los cientos de normas tenían sentido por separado. Pero, en su compleja totalidad, parecían algo absurdo para el propósito de Facebook de meter con calzador en diagramas dicotómicos los matices del discurso humano, la política y las relaciones sociales.

Un reglamento que Jacob consultaba a menudo —sobre cómo determinar si algo constituía un discurso de odio— tenía doscientas páginas repletas de argot. Pedía a los moderadores que hiciesen una especie de álgebra lingüística. Comparad la publicación comprobando un montón de listas: clases protegidas, insultos prohibidos, «comparaciones deshumanizadoras». Clasificad la gravedad de la publicación en una de las tres gradaciones previstas. Si la publicación hace referencia a un grupo, consultad una lista para comprobar si el grupo está prohibido, luego otra para ver si alguna de las palabras de

la publicación indica elogio o apoyo, luego otra para encontrar excepciones.

Y se pretendía que hicieran todo eso en entre ocho y diez segundos, trabajando de memoria. Con este tiempo limitado ni siquiera se podían mencionar todos los documentos relevantes, y menos aún revisar las normas que contenían. Y, desde luego, no era suficiente para pensar con detenimiento cómo aplicarlas de forma justa y segura. De hecho, no se animaba mucho a pensar. «Ese es el peor pecado, cuando te acusan de imponer tu propio juicio —dijo Jacob—. Nuestro trabajo consiste solo en seguir lo que dice el cliente.»

Yo le dije que entendía la posición de Facebook. Contrataba a miles de trabajadores en decenas de países. Sería poco o nada práctico y tal vez insensato darles voz a todos en la creación de esas normas.

Jacob asintió con la cabeza. Él solo quería que Facebook conociera los errores, las lagunas que había en las normas y los peligros que su equipo había visto en la plataforma pero que no podía eliminar hasta que eso apareciese por escrito en las guías oficiales. A fin de cuentas, dijo, los moderadores como él eran quienes veían la mayor cantidad de contenido. Pero sus jefes los disuadían a él y a sus compañeros de decir lo que pensaban. Les decían que, si las normas parecían equivocadas o incompletas, se callaran y no armaran líos. «Solo les interesa la productividad», dijo. Algunos de sus compañeros dejaron el trabajo. La mayoría, como necesitaban el trabajo, simplemente se quedaron y se callaron. «La gente se rinde», dijo Jacob. Sabían que, en cierto modo, la moderación de contenido era una misión condenada al fracaso. Ninguna guía podía contener el odio y la desinformación que los sistemas de Facebook estaban diseñados para producir en masa, aunque no fuera a propósito. Era como poner cada vez más ambientadores en los alrededores de una fábrica de residuos tóxicos mientras, al mismo tiempo, dentro la producción se incrementaba.

En cualquier caso, Facebook no parecía querer sus comentarios. Imperaba una mentalidad de programador: concebían las normas como líneas de código que unos moderadores intercambiables ejecutarían a modo de automatismo. En teoría, eso permitiría a Facebook aumentar la moderación al chascar los dedos. Y a eso mismo contribuiría subcontratar el trabajo a multinacionales que podrían atraer todos los empleados que fueran necesarios. Esa era la única

forma de mantener el ritmo de la expansión permanente de Facebook, que avanzaba demasiado deprisa para formar a revisores de contenido que reflexionasen por su cuenta.

Más adelante descubrí algo revelador. Distintos departamentos dentro de Facebook creaban guías usando estilos diferentes y a veces mentalidades distintas, lo que hacía que sus guías se contradijeran. Otras ni siquiera venían de Facebook, sino que las elaboraban agencias subcontratadas, sin el conocimiento de Facebook, para intentar mejorar los índices de cumplimiento dando a los moderadores unas reglas más estrictas que seguir. Algunas pretendían ser manuales de formación, pero, en lugar de eso, se empleaban como materiales de consulta. Parecía que la compañía no había asumido el papel de árbitro mundial, sino que se había visto inmersa en ese rol crisis a crisis, regla a regla. Sus directivos eran unos líderes supremos reticentes, recelosos ante posibles respuestas negativas, reacios a tomar decisiones, desempeñando su función en buena medida desde la sombra.

De vuelta a casa, limpié los archivos de cualquier huella digital y empecé a distribuirlas con discreción a analistas y expertos digitales de varios países. En India, Chinmayi Arun, experta en derecho, identificó errores preocupantes en las directrices para su país. Las directrices ordenaban a los moderadores que cualquier publicación que menospreciase una religión concreta infringía la legislación india y por tanto debía señalarse para ser eliminada. Era una restricción significativa al discurso de los 300 millones de usuarios de Facebook indios; y, además, era incorrecta. No existe una ley así.

¿Por qué Facebook, la compañía de la libertad de expresión, había impuesto una restricción tan estricta en la mayor democracia del mundo? Quizás fue por error. Pero en los archivos aparecían otros errores similares. Y luego esta pista: una serie de directrices de Pakistán alertaban a los moderadores de que no generasen un «incendio de relaciones públicas» tomando cualquier medida que pudiera «tener un impacto negativo en la reputación de Facebook o siquiera poner en riesgo jurídico a la compañía».

Otra diapositiva decía que la legislación india prohibía reclamar la independencia de Cachemira, una región que vive bajo una brutal ocupación militar india. Daba instrucciones a los moderadores de «estar atentos a» la expresión «Cachemira libre». Todo eso, dijo Arun,

en realidad era legal en India. Facebook me contó después que eso no pretendía prohibir a los usuarios expresar su apoyo por los derechos de los cachemires, sino solo animar a los moderadores a fijarse con mayor detenimiento en cualquier publicación que lo hiciera. Aun así, no estaba claro que esa distinción fuera evidente para los moderadores, a quienes se alertaba en distintos puntos de que la aplicación de las normas con demasiada laxitud o la aparición de escándalos conllevaría el bloqueo de Facebook en India. Pese a las promesas de la compañía de hacer una revolución de la libertad de expresión, a veces parecía prevalecer la deferencia ante las sensibilidades políticas de ciertos Gobiernos.

Errores había por todas partes. Las directrices para los Balcanes —un área de Europa afectada durante largo tiempo por conflictos entre milicias etnonacionalistas— estaban llenas de erratas, información desactualizada y errores extraños. A Ratko Mladić, un criminal de guerra aclamado todavía por los extremistas —el asesino de Christchurch había elogiado a Mladić en internet— se le identificaba erróneamente como «Rodney Young».

Facebook ejercía el control —en lo que Jacob destacaba una y otra vez como el defecto fatal del sistema— por medio de empresas subcontratadas sedientas de beneficios que no compartían ni sus objetivos ni sus valores. Y puso un ejemplo. En público, Facebook insistía en que a los moderadores no se les exigía que cumplieran unas cuotas determinadas de publicaciones revisadas o de velocidad. Las cuotas incentivarían la velocidad por encima de la precaución o la atención, lo cual forzaría a los moderadores a comprometer la seguridad de las comunidades que supervisaban. Pero Jacob me enseñó páginas de su oficina en las que aparecían cuotas de forma manifiesta: publicaciones que revisar, tiempo por publicación, contadores de cuántas publicaciones se habían revisado. Había documentos que demostraban que su salario se fijaba basándose en el ritmo de revisión. Las cuotas gobernaban sus vidas, impuestas por la agencia subcontratada con tal de que Facebook les encargase más trabajo. Eso se hacía sin el conocimiento explícito de Facebook y en contra de sus políticas, aunque también era un resultado previsible del hecho de que Facebook exigiera a las empresas externas que controlasen el coste, siempre al alza, de revisar miles de millones de publicaciones al día.

Facebook también afirmaba que los moderadores tenían acceso a servicios de salud mental. Eso había sido una respuesta a informaciones según las cuales había moderadores que desarrollaban trastornos de estrés postraumático al ver repetidamente pornografía gore y repugnante. Pero en su oficina eso nunca se había materializado, dijo Jacob. Muchos trabajadores terminaban agotados tras tan solo algunos meses, y a menudo volvían a trabajar en centros de atención al cliente. Eso era una señal de la desconexión entre Facebook y sus agencias, cuyos incentivos son únicamente tener unos costos bajos y una productividad alta. Cuantos más elementos conocía sobre cómo Facebook supervisaba a las empresas subcontratadas, menos sorprendente me parecía que el odio inundara un país tras otro sin que nadie lo revisara.

A los moderadores se les hacen auditorías de forma regular para comprobar su «precisión», el indicador que utiliza Facebook para determinar con qué frecuencia toman decisiones sobre el contenido en el sentido en que la compañía quiere que lo hagan. Una diminuta parte de esas auditorías las llevan a cabo empleados de Facebook. Pero la mayor parte la hacen otros trabajadores del mismo centro subcontratado. Las agencias se controlan a sí mismas. Su incentivo es revisar publicaciones lo más deprisa posible y luego declarar que alcanzan altos niveles de precisión.

La reducción de costes parece que se aplica sin escrúpulos. En la agencia de Jacob, si los moderadores encontraban una publicación en un idioma que nadie del equipo sabía leer, tenían la orden de marcarlo como aprobado, aunque los usuarios la hubiesen señalado como un discurso de odio peligroso. Se trataba de una revelación inquietante. No solo era que aquellas organizaciones de supervisión de Sri Lanka y Birmania tuvieran razón al afirmar que Facebook protegía activamente incitaciones directas al genocidio, sino que en algunas empresas subcontratadas tenían la política de hacerlo así.

El ejecutivo de Facebook que supervisaba las agencias admitió, al preguntarle sobre esta cuestión, que a veces la compañía tenía dificultades para controlar a las empresas externas, y declaró que esas prácticas eran una violación de las normativas de Facebook.³ Pero Facebook depende de la capacidad que tienen esas empresas de contratar y ampliar la plantilla con rapidez para sostener la expansión internacional que reporta a la compañía sus ingresos multimillonarios.

Aunque las empresas desobedezcan y engañen a Facebook, sus contratos tienden a aumentar. A fin de cuentas, los únicos que se resienten son los usuarios.

DUEÑOS DEL UNIVERSO

Hay muchos elementos que hacen que la tecnodirección de las redes sociales sea peculiar. La soberbia tanto de sus dimensiones como de su hermetismo. La creencia en que la política y los problemas sociales son problemas de ingeniería. La fe en los ingenieros para solucionarlos. La ingenuidad de pensar que los habían resuelto o, al menos, habían alcanzado un grado suficiente para seguir expandiéndose. Todo eso se remontaba, como muchas de las características definitorias de las plataformas, a los imperativos del capitalismo de Silicon Valley. O, mejor dicho, a un cambio drástico y reciente en ese modelo.

A finales de la década del año 2000, la relación de poder entre dos de las clases más importantes de Silicon Valley —los inversores y los creadores de empresas emergentes— se había invertido de golpe. Desde los primeros talleres de transistores, el poder había estado en manos de los inversores. Los empresarios incipientes necesitaban mucho dinero para fabricar su primer artilugio, más para llegar a su primer cliente y más aún para generar beneficios. Los inversores, para cubrir esos costes y asegurarse de que sus apuestas salían a cuenta, imponían una enorme supervisión.

«El pacto habitual era que el inversor de capital riesgo venía contigo a la empresa: se empleaba siempre el término “supervisión por parte de un adulto”», me explicó Leslie Berlin, historiadora de la Universidad de Stanford. Los inversores elegían a altos ejecutivos, al consejo de administración e incluso a un director ejecutivo con experiencia para que controlase al creador en su propia empresa. Cuando John Doerr invirtió 12,5 millones de dólares en Google, creada y dirigida por dos estudiantes de posgrado, les puso como jefe a Eric Schmidt, un veterano ejecutivo que les sacaba veinte años.

La era de la computación en la nube lo cambió todo. Ahora que cualquier empresario podía poner en marcha un negocio en internet y conseguir sus primeros mil clientes por su cuenta, y ahora que los

inversores de capital riesgo buscaban empresas emergentes de alto riesgo, de crecimiento trepidante y combustión rápida, necesitaban hacer más apuestas y, a su vez, los creadores de empresas los necesitaban menos a ellos. De repente, los inversores competían para convencer empresarios, y no al revés.

«Ahora, cuando los inversores en capital riesgo compiten para financiar a alguien, en lugar de que los emprendedores mendiguen dinero de rodillas, la dinámica de poder es totalmente distinta», pues los creadores de *start-ups* son quienes marcan las pautas cada vez más, me dijo Berlin. Una forma segura para que las apuestas de los inversores fueran más atractivas: se acabó la supervisión por parte de un adulto. Se acabaron los directores ejecutivos colocados como superiores de los empresarios. En el caso de Facebook, adiós al consejo independiente. «Aquello fue un cambio radical en el concepto de lo que se suponía que debía hacer un inversor de capital riesgo», dijo Berlin.

Esa evolución dio paso a la época del director ejecutivo-*hacker*-empresario de veintidós años que viste una sudadera con capucha. A la época de los equipos ejecutivos formados por los amigos del empresario. De las culturas empresariales que desafiaban con orgullo y osadía a los adultos con traje. «¿Por qué la mayoría de los maestros del ajedrez tienen menos de treinta años? —preguntó una vez Zuckerberg a un auditorio de 650 potenciales empresarios en la incubadora de *start-ups* de Paul Graham—. Pues porque los jóvenes son más inteligentes.»⁴ Así pues, animó a los potenciales ejecutivos del público a contratar a sus trabajadores siguiendo esa idea.

El propio Graham alentó a otros inversores a decantarse por la juventud, y dijo en una conferencia pública que «en el mundo del *software*, lo mejor es invertir en estudiantes, no en profesores».⁵ La experiencia empresarial, la experiencia vital y las titulaciones superiores eran lastres, no activos. Invertir en personas jóvenes, añadió, también tendía a ser más barato. El año anterior, la edad media de los creadores de empresas de su programa de *start-ups* había sido de veintitrés años. El resultado, dijo Berlin, era que recaía «mucho poder en las manos de personas muy jóvenes que piensan que la gente de su alrededor no tiene nada que ofrecer».

Muchos surgieron en el *boom* de las puntocom, una época de chefs de sushi que cocinaban *in situ* y masajistas. Las ventajas tenían la

intención de resolver un problema laboral.⁶ California prohíbe las cláusulas anticompetencia, lo que significa que los empleados pueden cambiar de trabajo al instante. Las empresas ofrecen oficinas lujosas y servicios de mayordomo para que los empleados no se vayan. Pero el estilo de vida puede confundir. Si todos vivimos como reyes y presidentes, entonces se supone que somos tan importantes como ellos, ¿no?

Las culturas empresariales, al absorber este espíritu, adoptaron una consigna muy ideológica: si creas artilugios para nosotros, no solo estás ganando dinero, estás salvando el mundo. La carrera armamentística entre misiones empresariales fue intensificándose hasta que, a finales de la década del año 2000, casi todas las empresas de la ciudad contaban a sus empleados que diseñar aplicaciones los colocaba, en términos de importancia, más o menos entre las Naciones Unidas y la Legión de Superhéroes. Era una cultura en la que tomar las riendas de la gobernación mundial no parecía muy descabellado.

Los egos se engrandecieron más aún con una práctica contable de la era digital: pagar a los primeros empleados con participaciones. *Start-ups* a medio hacer, a las que se pedía que convirtieran una inversión inicial de 100.000 dólares en una base de millones de consumidores, ofrecían, en lugar de unos salarios competitivos, lo que eran en realidad pagarés. Si la compañía se iba al traste, como les pasaba a la mayoría, esas participaciones no valían nada. Si la compañía tenía éxito —lo que podía significar una venta a Oracle por 300 millones de dólares—, esos primeros empleados podían jubilarse, siendo millonarios, a los veintiséis años. Para quienes sacaban el premio gordo, el dinero se veía como una afirmación de lo que inversores y empresarios siempre les habían dicho: sois distintos a los demás. Más inteligentes y mejores. Los legítimos dueños del universo.

La nueva era dio paso también a una clase distinta de inversores. Los inversores de capital riesgo apoyados por instituciones —como John Doerr actuando en nombre de Kleiner Perkins— «dedicaron varios años a hacer recortes después del estallido de la burbuja de las puntocom», afirma en sus memorias Roger McNamee, uno de los primeros inversores de Facebook. «Se extinguieron los inversores de proximidad, que eran personas individuales, sobre todo antiguos emprendedores y ejecutivos»,⁷ escribió McNamee. Personas ricas conocedoras del sector, que ponían su propio dinero, sin una empresa

que controlase los gastos ni unos accionistas a los que contentar. Tanto inversor como empresario ahora tenían mucha más libertad para hacer lo que quisiesen. Se estaba acabando la era de los guardianes y los guardarraíles.

Algunos de los primeros y más influyentes inversores de proximidad fueron los cofundadores de PayPal, entre ellos Peter Thiel, el ultraconservador que había dicho que el valor de la diversidad era un «mito» y que los *hackers* contestatarios «con ineptitud social como si tuvieran Asperger» eran los mejores jefes para una *start-up*. «El impacto de esos hombres transformó Silicon Valley», escribió McNamee, y su dinero financió buena parte de la era de las redes sociales. Pero el «sistema de valores» que trajeron consigo, según McNamee, «es posible que contribuyera a la ceguera de las plataformas de internet ante los daños derivados de su éxito».⁸

La política de los fundadores de PayPal tenía inclinaciones muy libertaristas: eran darwinistas sociales, desconfiaban del Gobierno y estaban convencidos de que las empresas sabían mejor qué hacer. Thiel llevó eso a tales extremos que en 2009 anunció: «Ya no creo que la libertad y la democracia sean compatibles».⁹ La sociedad ya no podía confiarse al «irreflexivo *demos* que guía a la llamada socialdemocracia», escribió, empleando el término griego para referirse a los ciudadanos. Solo las «empresas como Facebook» podían garantizar la libertad. Y solo si se deshacían de los grilletes de «la política», término con el que parecía referirse a la regulación, la responsabilidad pública y posiblemente la ley.

Financiaba proyectos pensados para crear ciudades flotantes dirigidas por empresas y la colonización del espacio, todo ello fuera de la jurisdicción de cualquier gobierno.¹⁰ Esas fantasías de ciencia ficción no hacían más que exagerar una vieja idea de Silicon Valley. Que los ingenieros y las empresas emergentes sabían mejor qué había que hacer. Era responsabilidad suya acabar con el *statu quo* e instalar una tecnoutopía en su lugar, como habían profetizado los redactores del manifiesto de los años noventa. Si los Gobiernos o los periodistas ponían objeciones, era solo porque los agentes tradicionales se aferraban a una autoridad que ya no les pertenecía.¹¹

Ese sentimiento de misión divina impulsaba a los inversores de proximidad de la Generación PayPal que seleccionaban las *start-ups* y los empresarios para rehacer el mundo en torno a su proyecto. Lo

llamaban desbancar a los poderes tradicionales. Uber y Lyft no solo ofrecerían una nueva forma de llamar un taxi, sino que abolirían y sustituirían la forma en que se había hecho toda la vida. Airbnb transformaría las opciones de alojamiento a corto plazo. Las tres habían recibido inversiones de PayPal. Muchas otras pretendían hacer el mismo desplazamiento violento. Amazon y el comercio minorista; Napster y la música. Solo unos pocos, como Thiel, proponían en serio hacerle a la gobernación internacional lo que Uber le había hecho al transporte en coche. Pero, una vez que las redes sociales se vieron en ese papel, debió de parecer solo como una continuación de su posición legítima. De la creencia de que la sociedad es una serie de problemas de ingeniería que esperan para ser resueltos.

CONVERGENCIA

Dos años después de que la elección de Trump desconcertara a los estadounidenses, que empezaron a preguntarse por la influencia de las redes sociales en su política, Renée DiResta y cuatro expertos más estaban sentados delante de las cámaras de televisión y las tarimas de paneles de madera de la Comisión Selecta sobre Inteligencia del Senado preparados para dar respuestas. «Este problema es una de las amenazas que caracterizan nuestra generación»,¹² dijo DiResta a los dieciséis senadores que tenía delante.

La sesión estaba destinada, nominalmente, a abordar la intromisión digital de Rusia. Pero los investigadores del Congreso, como muchos otros, comenzaban a pensar que la incursión rusa, aunque era perniciosa, había puesto al descubierto un peligro más profundo. «No se trata de arbitrar la verdad ni es una cuestión de libertad de expresión», dijo DiResta. El problema era la amplificación algorítmica, los incentivos digitales que llevaban a los usuarios a difundir propaganda de forma involuntaria y la facilidad con que los actores malvados podían «sacar partido de todo el ecosistema informativo para crear una apariencia de consenso popular». Como DiResta llevaba haciendo desde hacía años, dirigió la atención de su público de Moscú a Silicon Valley. «La responsabilidad de la integridad del discurso público está en buena medida en las manos de plataformas sociales privadas», afirmó. Por el bien común, añadió,

hablando en nombre de su equipo, «creemos que hay que hacer rendir cuentas a las plataformas tecnológicas privadas».¹³

De entrada, se había mostrado reticente a comparecer ante el Senado. «Me parecía que eso era un balón de fútbol —me contó, en referencia a la intromisión rusa—. ¿Me van a decir a todo “vale, entendido”? ¿Voy a estar peleándome con ciertos senadores republicanos que en ese momento eran sucedáneos de Trump?» Pero una sesión a puerta cerrada que había hecho para la comisión había sido «de una gran profesionalidad», dijo. «Todo fue verificación de hechos.» Impresionada y sorprendida por la sinceridad de algunos senadores que ella consideraba ultrapartidistas, aceptó. «Quería dejar claro con todas mis fuerzas que yo creía que se trataba de un problema sistémico —me contó DiResta— y que tal vez Rusia era la mejor manipulando el sistema, pero que otros seguirían haciéndolo.»

Cuantas más gigas de los datos proporcionados por las plataformas analizaba su equipo, dijo DiResta, más segura estaba de «que no importaba tanto si se trataba de Rusia, los antivacunas o unos terroristas. Esa era simplemente la dinámica que iba tomando forma como consecuencia de ese sistema». Durante meses, se habían producido señales de una gran convergencia en lo que tiempo atrás se había llamado «el manual de estrategia ruso» pero parecía cada vez más que los usuarios y los grupos sencillamente seguían los incentivos y funciones de las redes sociales. Los límites se habían vuelto borrosos, quizás para bien, entre los grupos que difundían desinformación de un modo estratégico al estilo ruso y los usuarios que la provocaban de un modo orgánico. Los propagandistas habían dejado de ser necesarios; el sistema, temía DiResta, era el que hacía el trabajo de verdad.

En un intento de abordar las preocupaciones de la sociedad, Zuckerberg publicó un artículo, algunas semanas después de la comparecencia de DiResta. «Uno de los principales problemas a los que se enfrentan las redes sociales —escribió Zuckerberg— es que, si no se aplica ningún control, los usuarios participan mucho más al entrar en contacto con contenido sensacionalista y provocativo.»¹⁴ Acompañaba el artículo de un gráfico en el que se veía que la participación ascendía cuando el contenido de Facebook se volvía más extremo, e iba subiendo hasta que llegaba al límite que Facebook permitía. «Nuestras investigaciones hacen pensar que, con independencia de dónde pongamos los límites de lo permitido, a

medida que un elemento se acerca a esa línea, las personas interactuarán más con él de media», escribía. «A gran escala», añadía, este efecto «puede perjudicar la calidad del discurso público y conducir a la polarización».

Guillaume Chaslot, el expleado de YouTube, no salía de su asombro. Se había pasado años intentando demostrar lo que Zuckerberg acababa de admitir. El contenido más extremo genera más participación de los usuarios, se promueve más dentro de la plataforma, lo cual polariza a los usuarios. «Es una locura cuando lo piensas bien», dijo Chaslot.

Zuckerberg había contado un detalle revelador: en sus investigaciones internas habían descubierto que las personas interactúan más con contenido extremo «incluso cuando después nos dicen que no les gusta ese contenido».¹⁵ Dicho de otro modo, como varios expertos y personas conocedoras del sector preocupadas habían intentado por todos los medios demostrar durante años, los usuarios no actuaban con una voluntad consciente, sino por algo entre el impulso, la tentación y algunos empujoncitos impuestos por el sistema. La solución que proponía Zuckerberg era, ¿cómo no?, que se modificase más el algoritmo. Lo entrenarían para que identificara contenido que no estaba prohibido del todo y rebajarían su nivel de promoción: una especie de prohibición a medias. Pero la naturaleza básica del sistema permanecería intacta.

«No cambió gran cosa —dijo Chaslot—. Aún puedes ver que, si te fijas en las publicaciones más compartidas sobre política en Facebook, son muy divisivas, tanto si son de extrema derecha como de extrema izquierda. —Y añadió—: Si Facebook siempre incita a los usuarios hacia los extremos, no importa que se hayan eliminado los grupos más extremos, porque Facebook crea más.» Por eso para él no había un problema de moderación, sino de amplificación.

Y ese problema, a pesar de las modificaciones de Zuckerberg, estaba empeorando. La compañía había renovado su algoritmo a principios de 2018:¹⁶ el año anterior, la participación había caído, y la modificación pretendía que volviera a subir. La nueva versión promovería o suprimiría cada publicación basándose en una puntuación asignada de forma automática. Los «me gusta» valían un punto, pero los emoticonos de reacción —un corazón, una carita triste o una carita de enfado— valían cinco, lo que significaba que una

publicación más emotiva valdría cinco veces más.¹⁷ Los comentarios cortos recibirían quince puntos, las comparticiones y los comentarios largos treinta, con lo cual se recompensaba aquello que provocaba un debate largo y emocional.

La participación volvió a incrementarse enseguida. Pero los usuarios expresaban estar menos satisfechos con la experiencia. El tráfico a medios de comunicación desde el canal de noticias descendió. «La desinformación, la toxicidad y el contenido violento tienen una preponderancia desproporcionada entre las comparticiones»,¹⁸ alertó un informe interno sobre el cambio, y en otro se afirmaba que había «efectos secundarios nocivos en secciones importantes del contenido público, como la política y las noticias». ¹⁹ Los insultos y los rumores proliferaban aún más que antes y desplazaban las noticias o la moderación. Partidos políticos de toda Europa se quejaron en privado a la compañía de que la inclinación sensacionalista de sus algoritmos «los había forzado a escorarse hacia lo negativo en sus comunicaciones por Facebook, lo que los había llevado a posiciones políticas más extremas»,²⁰ según un comunicado interno.

Más o menos cuando se publicó el artículo de Zuckerberg, un equipo de economistas de las universidades de Stanford y de Nueva York hizo un experimento para comprobar, de la forma más directa y rigurosa que se ha hecho nunca, cómo el uso de Facebook altera las opiniones políticas de los usuarios.²¹ Seleccionaron a unos 1.700 usuarios y los dividieron en dos grupos. A los de un grupo les pidieron que desactivasen su cuenta de Facebook durante cuatro semanas. A los otros no. Los economistas, empleando sofisticados métodos de estudio, monitorizaron el estado de ánimo diario de cada participante, su consumo de noticias, la precisión de su conocimiento sobre la actualidad y sobre todo sus opiniones políticas.

Los cambios fueron drásticos. Las personas que desactivaron su Facebook estaban más alegres, más satisfechas con su vida y menos nerviosas. El cambio emocional equivalía a entre un 25 y un 40 % de los efectos de ir a terapia: una disminución asombrosa para una pausa de cuatro semanas. Cuatro de cada cinco decían, al término del estudio, que desactivar Facebook había sido algo positivo. Los que habían dejado la plataforma también dedicaban un 15 % menos de su tiempo a consumir noticias. En consecuencia, estaban menos

informados sobre la actualidad: el único efecto negativo. Pero buena parte del conocimiento que se habían perdido parecía ser de contenido polarizador; información empaquetada con tal de satisfacer los antagonismos tribales. En conjunto, escribieron los economistas, la desactivación «redujo de forma significativa la polarización de las opiniones sobre cuestiones políticas y la exposición a noticias polarizadoras». El grado de polarización de esos participantes disminuyó *casi la mitad* de la cantidad que la polarización del estadounidense medio había aumentado entre 1996 y 2018: justo el periodo en que se había producido la crisis de polarización que había puesto en peligro la democracia. Insisto, *casi la mitad*.

A lo largo de 2018, a medida que aumentaron las pruebas, empezaron a aparecer medidas. Ese año, Alemania decretó que las plataformas de redes sociales eliminasen cualquier discurso de odio antes de veinticuatro horas después de haberse marcado como tal o se les impondrían multas. Australia anunció una investigación para introducir regulaciones «por primera vez en el mundo» sobre los perjuicios de las redes sociales, lo que calificó como un «punto de inflexión» en un momento en que «se está reconociendo a escala internacional que internet no puede ser ese sitio en el que los estándares comunitarios y el imperio de la ley no se cumplen».²² La Unión Europea impuso una serie de multas que superaban los mil millones de dólares a Google por abusos antimonopolio y amenazó con introducir regulaciones contra Facebook por discursos de odio, influencia electoral y desinformación: todo el abanico. «No voy a esconder que me estoy impacientando bastante —declaró la comisaria de Justicia de la UE, Věra Jourová en una rueda de prensa en referencia a Facebook—. Si no vemos avances, van a llegar las sanciones. No quiero estar negociando eternamente.»²³

Incluso Silicon Valley empezaba a interiorizar las críticas.²⁴ En una encuesta interna a 29.000 empleados de Facebook hecha ese octubre,²⁵ se halló que el porcentaje de empleados que decían estar orgullosos de trabajar en Facebook había disminuido de un 87 a un 70 % en tan solo un año. El porcentaje de quienes creían que su empresa hacía del mundo un lugar mejor había caído de un 72 a un 53 %, y sobre si veían con optimismo el futuro de Facebook, de alrededor de un 85 % a poco más de un 50 %. «Cuando empecé a trabajar en Facebook en 2016, mi madre estaba orgullosa de mí —dijo una

exdirectora de producto de Facebook en declaraciones a la revista *Wired*—. Podía ir por todas partes con mi mochila de Facebook y la gente me paraba y me decía: “Qué guay que hayas trabajado para Facebook”. Pues eso se acabó. —Y añadió—: Ahora cuesta ir a casa de tus padres el Día de Acción de Gracias.»

CIBERDEMOCRACIA

Siempre que ha habido una época democrática, ha estado gobernada por unos guardianes del acceso. Los *establishments* de los partidos dictan los programas y seleccionan quién sale en las papeletas. Los *establishments* mediáticos controlan quién tiene cuota de pantalla y quién no, a quién se le presenta como aceptable y a quién no. Las organizaciones empresariales y los grupos de intereses reparten la financiación que permite ganar elecciones. Las redes sociales, entre otros factores, socavaron el poder de esos guardianes del acceso. Por un coste cero, los candidatos podían armar sus propios mensajes públicos, su organización, sus imperios de captación de fondos y de este modo podían prescindir de los guardianes. Los *establishments* siguen teniendo influencia, pero, para bien o para mal, su control sobre la democracia es historia.²⁶

Silicon Valley, por supuesto, esperaba hacer justo eso. Al permitir que los usuarios se expresasen de forma directa y masiva, «en lugar de hacerlo a través de unos intermediarios controlados por unos pocos elegidos —escribió Zuckerberg en 2012 en una carta a los inversores—, esperamos que los Gobiernos se vuelvan más sensibles a la opinión de los ciudadanos».²⁷ Pero, en la práctica, las redes sociales no abolieron los *establishments*, sino que los sustituyeron. Sus algoritmos e incentivos ahora actuaban como guardianes del acceso y determinaban quién tenía protagonismo y quién se hundía. Y eso lo hacían basándose no en la popularidad sino en la participación, lo que, como Chaslot había demostrado en su análisis de YouTube, llevaba a que candidatos marginales obtuvieran mejores resultados.

Algunas personas llamaron a esa época «ciberdemocracia».²⁸ En Francia, a finales de 2018, lo llamaron «chalecos amarillos». Todo empezó cuando durante ese verano una petición que exigía que se bajasen los precios de la gasolina se propagó por las redes sociales y se

convirtió, en octubre, en la base de un grupo de Facebook en rápido aumento que animaba a los conductores a bloquear carreteras de su zona. El debate era espontáneo, y cada usuario culpaba del aumento del precio de la gasolina al agravio que prefería, a un exogrupo o a una conspiración, lo cual daba a todo el mundo un motivo para manifestarse. Un día de noviembre planificado de antemano, decenas de miles de personas de todo el país sacaron los chalecos reflectantes de sus coches —la legislación francesa obliga a tener uno a mano— y bloquearon carreteras cercanas. El espectáculo atrajo miradas, lo que generó nuevos adeptos.

Desde el principio, los Chalecos Amarillos, como se bautizaron ellos mismos, se identificaron como un movimiento sin líderes, radicalmente horizontal. Sin lugar a dudas, las redes sociales lo habían hecho posible. Nunca había existido una forma escalable, gratuita y abierta a todo el mundo para organizarse. Pero las plataformas también aplicaban una mano invisible guiadora. Sus elementos promocionales arrastraban a los usuarios hacia los grupos de alta participación que fomentaban la actividad, y donde las publicaciones más emocionales llegaban a los puestos más destacados.²⁹Y a eso mismo contribuía la naturaleza participativa de las plataformas, que recompensaba a los usuarios que ejercían su pertenencia al grupo con publicaciones y fotos de su corte de carretera. Y era una identidad compartida, en la que cualquier francés con un chaleco de un euro podía sentir que formaba parte de algo grande y representativo.

Pronto los Chalecos Amarillos se unieron en torno a una causa más grande: rehacer la democracia francesa a su imagen y semejanza. Una «asamblea ciudadana» implementaría la voluntad popular de forma directa, sin la mediación de representantes o instituciones. Los asuntos se decidirían por referendo. Los votantes tendrían la potestad de revocar a sus representantes en cualquier momento. Hasta que eso no se les concediera, prometían cortar las carreteras todos los sábados.

Los Chalecos Amarillos también publicaron una cacofonía de reclamaciones políticas, muchas de las cuales eran contradictorias.³⁰Querían que se anulasen varios impuestos y que se acabase con el *sinhogarismo*. Había que reducir el número de alumnos en las clases de preescolar y tenía que ser gratuito aparcar en el centro de las ciudades. La deuda nacional tenía que «declararse ilegítima» y no debía pagarse. Querían que se prohibiese la entrada de refugiados

al país. Una segunda lista, añadida una semana después, exigía que Francia saliese de la Unión Europea y de la OTAN, que los impuestos se redujeran a la mitad y que se acabase con casi toda la inmigración. Era un programa que solo podía surgir de las redes sociales: maximalista, incoherente, basado en la identidad y sin las ataduras de unos responsables que podrían haber dado coherencia a la lista. Identidad en estado puro.

Los Chalecos Amarillos fueron, en algunos aspectos, un éxito histórico, el movimiento ciudadano más grande desde por lo menos la década de los sesenta en el país que prácticamente inventó tales movimientos. Pero tuvo poco impacto. Todos los partidos políticos intentaron sacar provecho de su energía. Y también los sindicatos. Y los estudiantes. Pero no llegaron a ninguna parte. Luego el movimiento fue apagándose, como si no hubiera existido nunca. Incluso más de siete años después de la Primavera Árabe y cualesquiera que fueran las lecciones que pudiesen haber extraído los activistas de internet, la nueva democracia de las redes sociales generaba mucho caos pero, extrañamente, pocos resultados.

Eso resultó que representaba una tendencia —reveladora— identificada por Erica Chenoweth, una experta de la Universidad Harvard sobre la resistencia civil. La frecuencia de los movimientos de protesta masivos iba en aumento en todo el mundo desde los años cincuenta, descubrió Chenoweth, y en los últimos tiempos se había acelerado.³¹ Entre la década del año 2000 y la de 2010, en promedio los episodios por año se habían disparado un 50 %. Su índice de éxito también se había incrementado, año tras año, durante décadas. En torno al año 2000, un 70 % de los movimientos de protesta que reclamaban cambios sistémicos alcanzaban éxitos. Pero luego, de repente, esa tendencia se invirtió. Los movimientos empezaron a fracasar, al tiempo que se volvían más frecuentes. Ahora, descubrió Chenoweth, solo terminaban con éxito un 30 % de los movimientos de masas. «Ha habido un auténtico cambio», me contó, y calificó el descenso de «asombroso». Casi todos los meses, en algún país estallaban protestas a escala nacional: en el Líbano por la corrupción, en India por la desigualdad de género, en España por el separatismo catalán. Muchas a una escala que excedía los movimientos más transformadores del siglo xx. Y la mayoría de ellos terminaban apagándose.

Para explicar eso, Chenoweth se basó en una observación de Zeynep Tufekci, la profesora de la Universidad de Carolina del Norte: las redes sociales facilitan a los activistas la organización de manifestaciones y la rápida captación de seguidores en unas cantidades antes impensables, pero en realidad eso podría ser un lastre.³² Las redes sociales, aunque de entrada se consideraron una fuerza que favorecía la liberación, «en realidad dan ventaja a la represión en la era digital mucho más que a la movilización», dijo Chenoweth. Los dictadores habían aprendido a aprovecharlas en su favor, empleando sus mejores recursos para inundar las plataformas de desinformación y propaganda.

Los efectos en las democracias eran más sutiles pero aun así potentes. Chenoweth citaba, a modo de comparación, el Comité Coordinador Estudiantil No Violento (SNCC, por sus siglas en inglés), una organización estudiantil de la época de los derechos civiles. Antes de las redes sociales, los activistas tenían que movilizar a la gente difundiendo su mensaje entre la comunidad y construyendo una organización. Se reunían casi a diario para formarse, diseñar estrategias y deliberar. Era una labor agotadora y de años. Pero hacía que el movimiento fuera duradero, construido sobre unos vínculos en el mundo real y con cadenas de comando. Permitía a movimientos como el SNCC perseverar cuando la situación se complicaba, responder de manera estratégica a hechos y traducir victorias en la calle en cambios políticos.

Las redes sociales hacen posible que las protestas se salten muchos de esos pasos y ponen más cuerpos en las calles más deprisa. «Eso puede dar a la gente una falsa sensación de confianza —dijo Chenoweth—, porque hay un compromiso menor.» Sin la infraestructura subyacente, los movimientos hechos en las redes sociales son menos capaces de organizar unas reclamaciones coherentes, coordinarse o actuar siguiendo una estrategia. Y alejar la energía popular de la forma más exigente de organización impide que surjan movimientos tradicionales. Eso era lo que había prometido Zuckerberg: movimientos ciudadanos más grandes y sin líderes. Pero Facebook, como los demás gigantes de las redes sociales, se había convertido en una institución consolidada por derecho propio. Y al igual que la mayor parte de los guardianes del acceso, tendía a proteger al *establishment* y el *statu quo* del que dependía para mantener

DIRECTORES EJECUTIVOS CON ACTITUD BÉLICA

En mayo de 2016, el portal de información tecnológica *Gizmodo* publicó la que probablemente sea todavía una de las noticias de mayor calado que se hayan escrito sobre las redes sociales, con el titular «Extrabajadores de Facebook: de forma habitual eliminábamos noticias conservadoras». ³³El artículo trataba sobre un pequeño *widget* que había en la página de inicio de Facebook llamado «tendencias», que mostraba varios temas y titulares. No generaba mucho tráfico: un 1,5 % de los clics. Un algoritmo ayudaba a identificar temas de tendencia, pero la compañía, en una decisión inhabitual, cedió esas decisiones al juicio humano. Varios periodistas seleccionaban los temas y escribían textos breves describiéndolos.

Un extrabajador subcontratado de ese equipo contó a *Gizmodo* que el equipo se había constituido para eliminar temas de orientación conservadora. El extrabajador, que se identificaba como conservador, dijo que los editores de Facebook le habían pedido que no incluyera noticias virales sobre, por ejemplo, un extrabajador del Servicio de Impuestos Internos (IRS, por sus siglas en inglés) a quien los republicanos habían acusado de atacar a grupos conservadores. (Las pruebas que daba eran vagas y la noticia resultó ser falsa.) Los editores también le dijeron que no eligiese noticias procedentes de sitios de extrema derecha como *Breitbart* hasta que un medio de comunicación tradicional las hubiera confirmado. El artículo de *Gizmodo* presentaba esas decisiones como una prueba del sesgo anticonservador de Facebook. También acusaba a la compañía de ordenar a sus trabajadores externos que «manipularan de forma artificial el módulo de tendencias» con la «inyección» de noticias que aún no se habían hecho virales. Los dos ejemplos que daban eran la desaparición del vuelo 370 de Malaysia Airlines y los atentados terroristas contra el semanario *Charlie Hebdo*.

Este es un caso infrecuente en que Facebook hizo todo lo que más adelante expertos y organizaciones digitales le pedirían que hiciera. Imponer supervisión humana sobre sus algoritmos. Privilegiar la verdad y la credibilidad antes que la viralidad y la participación.

Comprobando de forma ocasional los peores impulsos de su sistema, incluso de sus usuarios. Pero el artículo, enmarcado de aquella forma, ofrecía a los republicanos una oportunidad para atizar unos agravios que transformaron la política en torno a las redes sociales y, en última instancia, las propias plataformas.

«Facebook tiene la capacidad de ejercer una gran influencia sobre las elecciones presidenciales —dijo el Comité Nacional Republicano en un comunicado de prensa, mezclando el canal de noticias de Facebook, mucho más potente, con su *widget* de tendencias, ignorado en buena medida—. Es extremadamente alarmante enterarse de que ese poder se está empleando para silenciar puntos de vista y noticias que no se corresponden con el programa político de alguien.»³⁴El senador republicano John Thune mandó una carta a Zuckerberg pidiéndole que la compañía diera explicaciones a los congresistas sobre ese asunto. Thune presidía la comisión de comercio, que supervisa la Comisión Federal de Comercio, que había estado investigando Facebook. La compañía, asustada, mandó a su equipo en Washington para calmar a Thune. También invitó a una veintena de destacados conservadores, entre ellos un asesor de la campaña de Trump, para conocer a Zuckerberg en persona. Y funcionó. Los asistentes se criticaron con mayor acritud entre ellos que a Facebook.³⁵Uno de ellos, Glenn Beck, redactó una nota sobre la reunión en la que elogiaba a Zuckerberg como un gran defensor de la libertad de expresión.

Facebook despidió a los trabajadores externos que supervisaban el *widget* de tendencias un viernes y dejó que los algoritmos asumieran el control absoluto. El lunes, la noticia que aparecía en primer lugar era un enlace a una noticia falsa en la que se acusaba a la presentadora de Fox News Megyn Kelly de apoyar a Hillary Clinton.³⁶El blog de extrema derecha que lo había publicado fue identificado más adelante como uno de los motores principales de noticias falsas en Facebook durante las elecciones de 2016.³⁷

Ese verano, en julio, tan solo unos días antes de que la Convención Nacional Republicana nominara formalmente a Donald Trump como candidato a la presidencia, Zuckerberg fue interpelado en un congreso financiero por Rupert Murdoch y Robert Thomson, creador y director ejecutivo, respectivamente, de News Corp, empresa propietaria de Fox News y muchos otros medios conservadores a

escala internacional.³⁸Facebook, le dijeron, se estaba comiendo el negocio de la información al usar su contenido para hacerse con los espectadores y los ingresos publicitarios de los medios. Los cambios no anunciados a los algoritmos, le dijeron también, ponían en peligro la capacidad de News Corp de mantener el tráfico. Amenazaron con que News Corp haría presión pública ante Gobiernos y organismos reguladores para que evaluaran si el poder de mercado de Facebook estaba infringiendo las normas antimonopolio.

No amenazaron con utilizar sus medios de comunicación contra Facebook. Pero no habría sido insensato que Zuckerberg así se lo temiera. En los siguientes meses, Fox News habló de forma clara y directa sobre el supuesto sesgo anticonservador de Facebook.³⁹Calificó la campaña de Facebook «get out the vote» (‘saca el voto’) como una estratagema para aumentar la participación demócrata. Arremetió contra Facebook por «eliminar» lo que ellos llamaban «noticias destacadas», como el percance de salud de Hillary Clinton (de hecho, la plataforma había eliminado varias publicaciones con noticias falsas) y por «censurar» a la «activista» Pamela Geller, una figura de la extrema derecha a quien le eliminaron la página de Facebook tras difundir conspiraciones racistas.

Zuckerberg intentó apaciguar a los conservadores señalando, en comunicados y publicaciones a lo largo de 2016, que Trump tenía más seguidores que cualquier otro candidato presidencial y que Fox News tenía más interacciones en su página que cualquier otro medio. («[Los demás] ni se acercan»,⁴⁰dijo Zuckerberg.) La relación de la empresa con la campaña de Trump, que destinaba grandes sumas a anuncios en Facebook, se mantuvo cordial.

Después de las elecciones, a medida que aumentaban los indicios de que las grandes plataformas de redes sociales habían promocionado contenido falso y polarizador que había favorecido a Trump —parte de él con el apoyo de Rusia—, los republicanos vieron con claridad que su victoria cada vez estaba más empañada o quizás, por los indicios de coordinación indirecta entre la campaña de Trump y agentes rusos, incluso deslegitimada. Trataron de invertir el relato. Las redes sociales no habían favorecido a los republicanos, afirmaban, sino que los habían silenciado. Los republicanos, con el control de las dos cámaras del Congreso, queriendo distraer la atención de la investigación sobre la influencia rusa, organizaron comparecencias en

las que se daba el mensaje contrario y se afirmaba que los liberales de Silicon Valley eran la verdadera amenaza para la democracia.

Los republicanos encontraron la prueba del delito perfecta en un artículo de *Vice News* publicado en 2018 titulado «Twitter está “bloqueando de forma subrepticia” a destacados republicanos».⁴¹ El periodista había descubierto que el cuadro de búsqueda de Twitter no ofrecía automáticamente ciertas cuentas conservadoras cuando él introducía las primeras letras de sus nombres. La noticia parecía deberse a un malentendido técnico, pero los republicanos aprovecharon su oportunidad.⁴² En un tuit, el representante por Florida Matt Gaetz dio a entender que Twitter lo castigaba por haber interrogado al director ejecutivo de la compañía en una comparecencia reciente.⁴³ La presidenta del Comité Nacional Republicano dijo que la compañía «reprime las voces conservadoras»,⁴⁴ mientras que Donald Trump Jr. tuiteó: «Ya está bien de esa mierda»,⁴⁵ exigiéndole al director ejecutivo de Twitter «#StopTheBias» (‘Acaba con el sesgo’). Su tuit obtuvo 16.500 interacciones; el de Gaetz 20.400. No es una fuerte represión. Luego Trump padre también tuiteó «#StopTheBias», con lo que inició una oleada de mensajes conservadores que duró meses sobre la tiranía de las grandes tecnológicas.

Durante su presidencia, Trump y su partido pintaron las redes sociales como si fueran unos agentes antirrepublicanos. En reiteradas ocasiones amenazaron con investigar, regular o incluso fragmentar las compañías, amenazas que tuvieron una enorme fuerza para motivar a Silicon Valley. Los temores ante la aplicación de leyes antimonopolio —multas como forma de castigo o regulaciones como medio de contención ante prácticas monopolísticas— llevaban tiempo cerniéndose sobre el sector. En 1969, el Departamento de Justicia había iniciado un plan para fragmentar IBM, que controlaba el 70 % del mercado de los ordenadores.⁴⁶ El juicio se alargó durante trece años, tiempo en que IBM, para evitar que los organismos reguladores pudieran incriminar a la empresa, restringió su actividad empresarial. Cuando el Gobierno desestimó las acusaciones en 1981, la cuota de mercado y los ingresos de IBM se habían reducido en tal grado que un titular del *New York Times* proclamó: «Se acabó el dominio».⁴⁷

Pero lo que más aterrorizaba a las grandes tecnológicas era el caso de Microsoft. Una batalla que duró años con los organismos

reguladores había culminado, a mediados de los años noventa, con una sentencia judicial que obligaba a la empresa a dividirse en dos. (El Departamento de Justicia había acusado a Microsoft de explotar su hegemonía en un mercado para monopolizar otros.) La sentencia fue revocada en apelación —el juez había hablado sobre el caso con periodistas durante el proceso de deliberación, lo cual ponía en duda su imparcialidad— y la Administración entrante, con Bush a la cabeza, desestimó el caso. No obstante, el precio de las acciones de Microsoft se había reducido a la mitad; se había interrumpido su entrada a los servicios de internet, de lo que nunca se recuperaría; y su posición ante la sociedad y los organismos reguladores estaba tan deteriorada que Bill Gates, su creador, dimitió. Años después, aconsejaría a Zuckerberg que no repitiese lo que él veía como un error: contrariar a Washington e ignorar a los políticos que él creía que estaban equivocados. «Le dije: “Abre unas oficinas allí, ahora mismo” —recordaba Gates, en referencia a Washington, donde Facebook y Google empezaron a destinar millones a labores de influencia—. Y Mark lo hizo, y me debe una.»⁴⁸

Se acumulaban las multas regulatorias: 22 millones de dólares a Google en 2012, 170 millones de dólares en 2019. Cien millones a Facebook en 2019, luego unos 5.000 millones de dólares de récord también ese año, impuestos por la Comisión Federal de Comercio por violaciones de la privacidad de los usuarios. Se hablaba incluso de fragmentaciones forzosas. El senador Richard Blumenthal, un demócrata de Connecticut, sostuvo en un artículo de 2018 que el caso de Microsoft proporcionaba un modelo explícito para abordar la situación de Facebook, Google y Amazon.⁴⁹

Defenderse de esos ataques requeriría algo más que hacer labores de influencia. A partir de ese año, Zuckerberg y otros directores de empresas tecnológicas adoptaron una postura que el inversor de capital riesgo Ben Horowitz ha denominado como la de un «director ejecutivo con actitud bélica». Si los organismos reguladores no se aclaran o el mercado se hunde, había escrito Horowitz en una entrada de blog, las empresas necesitan un líder que «se salte el protocolo para ganar», que «utilice los insultos con propósito», que «sea completamente intolerante» con los empleados que se alejen de la estrategia corporativa, que «ni permita la creación de consensos ni tolere los desacuerdos» y que avergüence a los trabajadores delante de

sus compañeros para dejar claro un argumento.⁵⁰

Horowitz atribuía esa estrategia a su héroe Andy Grove, el antiguo director de Intel. Él había adoptado el término «director ejecutivo con actitud bélica» de una frase de *El Padrino* (Tom Hagen: «Mike, ¿por qué estoy fuera?»; Michael Corleone: «No eres un *consigliere* con actitud bélica»), que reprodujo en su publicación junto con una letra de una canción de rap, como parte de su campaña para reinventar a los multimillonarios tecnológicos como la nueva contracultura: unos genios con carácter fuerte vestidos con jerséis de cuello de pico. Su prestigio le venía de Andreessen Horowitz, una empresa de inversión que habían creado en 2009 él y el fundador de Netscape Marc Andreessen. (El mentor de Andreessen, John Doerr, había sido discípulo de Grove, un ejemplo más de la insularidad distorsionadora de Silicon Valley.)

Si William Shockley había encarnado la era de los semiconductores en el sector, Andy Grove la de los microchips y Peter Thiel la era de los inicios de la web, Andreessen y Horowitz personificaban la era de las redes sociales. No solo por las empresas en las que habían invertido (Facebook, donde Andreessen estaba en el consejo de administración, además de Twitter, Slack, Pinterest, Airbnb, Lyft y Clubhouse), sino por cómo lo habían hecho. Institucionalizaron, como estrategia de inversión, la tendencia de Silicon Valley por los directores ejecutivos jóvenes, inexpertos y sin restricciones. Prometían elevar a «creadores de empresas técnicos» con poca experiencia o conocimientos más allá de la ingeniería y luego dejarlos libres de la supervisión de un adulto o de cualquier expectativa de conducta empresarial normal.

El concepto de director ejecutivo con actitud bélica convirtió ese arquetipo en una filosofía empresarial. Unos programadores que creaban una aplicación móvil o un comercio electrónico debían ejercer el poder exactamente como las estrellas del rock iconoclastas que creían ser. Si alguien se ofendía —empleados con la piel fina, hombres trajeados cuadriculados o la Comisión Federal de Comercio—, lo mejor era salir con todo y hacerles saber quién eras. Una de las charlas de Horowitz sobre cómo gestionar una empresa trazaba un extenso paralelismo entre los creadores de *start-ups* y el líder de la rebelión de los esclavos de Haití en 1791,⁵¹ quien había tenido éxito —decía Horowitz— forzando a los haitianos a superar la «cultura de la

esclavitud», lo mismo que los directores ejecutivos tecnológicos harían con los empleados y ciudadanos que estaban perdidos sin su liderazgo. Era la clase de soberbia que podía conducir a un desarrollador web que había dejado la universidad a llegar a la conclusión de que tenía lo necesario para dictar los términos de las relaciones sociales humanas por todo el mundo, y que cualquier persona lo bastante anticuada para cuestionar eso era un esclavo que no lo pillaba. Ser un director ejecutivo con actitud bélica también proporcionaba una especie de cobertura moral. Si había que destruir a los competidores, si los trabajadores ponían objeciones éticas o los medios te acusaban de ser cómplice de la destrucción del tejido social, era porque no entendían que estábamos en guerra.

En 2018, Zuckerberg leyó un libro de Horowitz en que se desarrollaba esa estrategia.⁵²Ese junio, reunió a los cincuenta ejecutivos de mayor rango de la empresa para anunciarles que Facebook estaba en guerra y que él ahora era un director ejecutivo con actitud bélica.⁵³Iba a tolerar menos disensión, exigiría más obediencia y se enfrentaría a los enemigos de Facebook. En una asamblea con el conjunto de los trabajadores, declaró que la cobertura informativa sobre las violaciones de privacidad de Facebook, por las que Facebook tendría que pagar varias multas regulatorias, eran «sandeces». Echó la bronca a Sheryl Sandberg, la segunda ejecutiva de mayor rango de la compañía y su asesora durante mucho tiempo. Facebook contrató una empresa de relaciones públicas turbia, que sembró información negativa —en parte falsa— sobre personas que criticaban Facebook.⁵⁴

Destacados inversores de capital riesgo anunciaron que Silicon Valley estaba en guerra contra un sector mediático nacional deshonesto que quería castigarlos por su éxito.⁵⁵(«Lo pillamos: nos odiáis. Y sois nuestros competidores»,⁵⁶tuiteó una persona.) Algunos intentaron contraatacar y pidieron que se prohibiera a todo Silicon Valley colaborar con medios de comunicación o, en un caso, se ofrecieron a pagar en bitcoins a algunos usuarios para que acosaran a periodistas críticos por internet.⁵⁷

Unas semanas después de la declaración «bélica» de Zuckerberg, Facebook celebró una reunión en la que se planteó modificar su algoritmo para dar mayor prioridad a los medios de comunicación serios.⁵⁸Eso podía restablecer la confianza en Facebook, coincidieron

algunos ejecutivos. Pero en contra de la idea estaba el lobista Joel Kaplan, antiguo alto funcionario de la Administración Bush. Desde la elección de Trump, Kaplan parecía actuar, con el beneplácito de Zuckerberg, como el representante del Partido Republicano en Facebook, una labor que llevaba el título de vicepresidente de políticas públicas internacionales. Kaplan sostenía que el cambio provocaría acusaciones del Partido Republicano de que Facebook favorecía a los liberales, lo que en la práctica convertiría el punto de vista de Trump —que los periodistas de los medios tradicionales eran agentes demócratas— en la política corporativa de Facebook. Kaplan se impuso.

Ese mismo año, Kaplan logró que se archivase uno de los informes internos de la compañía en que se afirmaba que los algoritmos de la plataforma promovían contenidos divisivos y polarizadores.⁵⁹ Él y otros empleados objetaron que abordar el problema afectaría de un modo desproporcionado las páginas conservadoras, las cuales difundían un alto porcentaje de la desinformación. Era mejor dejar que los usuarios estuvieran mal informados. Esa no fue la última vez que se sacrificaría el interés público para evitar siquiera hipotéticas objeciones de los republicanos, aunque fuesen infundadas.

El cortejo de los republicanos por parte de Facebook, quienes entre 2018 y 2019 tuvieron el control de las palancas de la supervisión federal, fue abundante. La empresa contrató a Jon Kyl, un exsenador republicano, para que elaborase un informe sobre cualquier sesgo anticonservador que hubiera en la plataforma. El informe en buena medida reformulaba las acusaciones del «#StopTheBias» de Trump, lo que permitió a Facebook decir a los críticos del Partido Republicano que estaba estudiando el asunto y seguir las recomendaciones de Kyl. Zuckerberg asistió a cenas extraoficiales con influyentes conservadores,⁶⁰ como el presentador de Fox News Tucker Carlson, quien había acusado Facebook de querer «la muerte de la libertad de expresión en Estados Unidos».⁶¹ La plataforma contrató al *Daily Caller*, el portal de noticias de derechas creado por Carlson, para que participara en su programa de verificación de datos, con lo cual le confirió poder para determinar qué era verdad en la plataforma. Facebook anunció que permitiría que los políticos mintiesen en la plataforma y que les concedería una especial flexibilidad con respecto

a los discursos de odio, unas normas que parecían escritas para Trump y sus aliados.⁶²

«Llevaba menos de un año en Facebook cuando me vi inmersa en una consulta urgente: la campaña del presidente Trump se quejaba de una disminución en el número de visualizaciones —recordó en Twitter Sophie Zhang, experta en ciencia de datos de Facebook—; nunca me habían pedido que investigase algo parecido para nadie.»⁶³ Esa clase de conciliación con dirigentes políticos parecía ser una estrategia mundial. Entre 2018 y 2020, Zhang marcó decenas de incidentes de dirigentes extranjeros que difundían mentiras y odio para beneficio propio, pero —según ha contado— la desautorizaron de forma sistemática.⁶⁴ Cuando la despidieron, se negó a aceptar una indemnización de 64.000 dólares para no desacreditar a la empresa, pues quería publicar su memorando de salida, de 7.800 palabras, en el que relataba lo que consideraba una práctica deliberada para permitir que los políticos hiciesen un mal uso de la plataforma, también en países donde incluso podía haber violencia sectaria y el autoritarismo iba en aumento. «Sé que ahora ya tengo las manos manchadas de sangre»,⁶⁵ escribió.

En 2019, la dictadura comunista de Vietnam hizo llegar a Facebook un mensaje privado:⁶⁶ la plataforma tenía que censurar a los críticos con el Gobierno o el Gobierno vietnamita bloquearía la plataforma en el país. Zuckerberg aceptó, como desvelaron después algunos empleados, con lo cual Facebook —«la empresa revolucionaria», tal como él la había denominado— se convertía en secreto en una herramienta de la represión autoritaria. Aunque argumentó que para los ciudadanos vietnamitas era mejor acceder a un Facebook libre a medias que el bloqueo total de la plataforma, su secretismo inicial y el historial de Facebook despertaron dudas sobre la integridad de las intenciones de Zuckerberg. Una organización calculó que la presencia en Vietnam reporta a Facebook mil millones de dólares al año.⁶⁷

La compañía también anunció ese año que Facebook dejaría de revisar la veracidad o la precisión de los anuncios políticos.⁶⁸ Solo se eliminarían aquellos que infringiesen de forma extrema las normas, como los que llamasen a la violencia. Se presume que Trump, que había destinado cantidades generosas de dinero a Facebook antes de eso, fue el mayor beneficiario, así como cualquier persona como él.

Unos doscientos cincuenta empleados firmaron una carta abierta — una insólita demostración de disensión pública— en la que pedían a Zuckerberg que diera marcha atrás en esa política, que «haría aumentar la desconfianza en nuestra plataforma» y «echaría por tierra las labores para la integridad del producto»⁶⁹ que iban encaminadas a proteger elecciones.

Yaël Eisenstat, exasesora de la Casa Blanca que supervisó la política de Facebook ante citas electorales durante los seis meses previos a dejar la empresa, publicó una columna sobre esa polémica.⁷⁰ Afirmaba haber visto a responsables de la compañía trabajar duro para conciliar la integridad democrática con la misión de Facebook, pero que al final los habían desautorizado «las pocas voces que en última instancia deciden la orientación general de la compañía». Facebook, alertaba Eisenstat, estaba suspendiendo «el examen clave que va a determinar si algún día pondrá de veras a la sociedad y la democracia por delante de los beneficios y la ideología».

REPARAR UN AVIÓN EN PLENO VUELO

Cada dos jueves por la mañana, varias decenas de empleados de Facebook se pasean por una sala de reuniones de cristal para participar en lo que de entrada se parece a cualquier otra reunión de plantilla de Silicon Valley. Eligen lo que quieren de un desayuno de *catering*, dejan caer botellas de agua *eco-friendly* y jueguean con el programa de videoconferencia. Luego se acomodan y debaten, en una horita, cómo modificar y ajustar el experimento constante de la compañía sobre la gestión de las relaciones sociales y el discurso político en el mundo.

La semana en que yo estuve de visita, en octubre de 2018, un tema importante eran los estándares para la venta de animales en las páginas de comercio de la plataforma, al estilo de Craigslist. Los trabajadores debatían cómo equilibrar los derechos e intereses de la fauna, los animales domésticos, el ganado, los ganaderos y los criadores de animales. Se revisarían las regulaciones europeas. Se consultaría a un grupo externo. Alguien planteó, a modo de consideración, la importancia económica del comercio de ganado para los usuarios de África. Sus decisiones se traducirían en las normas y

directrices maquinales que se daban a Jacob y miles de moderadores más de todo el planeta, quienes a su vez las usarían para influenciar a miles de millones de personas.

Facebook me había invitado a visitar la empresa al enterarse de que había tenido acceso a sus reglamentos internos: los que me había facilitado Jacob. Querían que entendiera los archivos en contexto y, como me dijo un empleado, querían empezar a abrir sus procedimientos un poquito al mundo exterior. Monika Bickert, que dirige el equipo de políticas, me contó, en una de la docena de entrevistas que la compañía había organizado, que el proceso de elaboración de normas solía ser reactivo. Un trabajador, una noticia o un grupo de interés podían identificar un problema. Alguien redactaba una nueva norma, que se debatía en la reunión y luego pasaba por un proceso de revisión. Podía probarse en un mercado concreto antes de desplegarse a otros lugares. «No trazamos esas líneas en el vacío», dijo Bickert al describir las normas y principios legales que los guiaban. Pero cuanto más complejas se volvían las reglas, más difícil era implementarlas por medio de moderadores como Jacob, sobre todo debido a los cortísimos plazos que se deban a los moderadores para tomar decisiones. Bickert era consciente de la contradicción. «Cada vez que introducimos matices en nuestras políticas, eso hace que nos cueste más aplicarlas de un modo sistemático y adecuado en todo el mundo —dijo—. Existe una verdadera tensión entre querer contar con unos matices que contemplen todas las situaciones y querer disponer de una serie de políticas que podamos aplicar con precisión y que podamos explicar con claridad.»

Yo me di cuenta de que basculaba entre la compasión y el escepticismo por los responsables de las políticas de Facebook. Eran considerados y, la mayoría de las veces, humildes con respecto a su capacidad de abordar las complejidades conductuales y las políticas que tenían la misión de gestionar. Algunos habían comenzado a trabajar para la compañía pensando que harían una labor más positiva mejorando Facebook desde dentro que criticando desde fuera. Y habían quedado atascados en la imposible tarea de actuar como limpiadores de los destrozos hechos por los equipos de crecimiento de la compañía, con más recursos y más aclamados. Mientras ellos se preocupaban por problemas como los discursos de odio contra los refugiados o la desinformación en elecciones sensibles, los ingenieros

de la sala de al lado maximizaban la participación de los usuarios de formas que, casi de manera inevitable, empeoraban esos problemas.

Aun así, era difícil separar la benevolencia de su trabajo del grado en que pretendía —como se afirmaba sin rodeos en algunos documentos sobre políticas— proteger a Facebook de los reveses públicos o de la regulación. Acabé imaginando al equipo de políticas de Facebook como algo parecido a los científicos de Philip Morris encargados de desarrollar un filtro mejor y más seguro. En cierto modo, reducir los carcinógenos ingeridos por miles de millones de fumadores de todo el mundo salvaba o prolongaba vidas a una escala que pocos de nosotros podrían hacerlo nunca. No obstante, desde otro punto de vista, esos científicos estaban trabajando para la compañía tabacalera y priorizaban la causa de la venta de unos cigarrillos que perjudicaban a una cantidad ingente de personas.

No me sorprendió, pues, que todas las personas con las que hablaba en Facebook, por más inteligentes o introspectivas que fueran, expresasen un total convencimiento en que el producto no era perjudicial de forma inherente. Que no había pruebas de que los algoritmos u otras características arrastrasen a los usuarios hacia el extremismo o el odio. Que la ciencia aún no había determinado si los cigarrillos eran adictivos y causaban cáncer realmente. Pero, del mismo modo en que resultó que Philip Morris disponía de montones de estudios que demostraban riesgos para la salud que sus ejecutivos insistían en que no existían, los propios investigadores de Facebook habían ido acumulando pruebas, en rimeros de informes y montones de experimentos internos, que llegaban a una conclusión que declararían explícitamente en agosto de 2019: «Los mecanismos de nuestra plataforma no son neutros».⁷¹

En un informe interno sobre el odio y la desinformación se habían encontrado, según escribían sus autores, «pruebas concluyentes de que los mecanismos de nuestro producto principal —como la viralidad, las recomendaciones y la optimización para conseguir una mayor participación— explican en un porcentaje significativo por qué esos tipos de discurso aparecen en la plataforma». El informe, filtrado más adelante a los medios y a la Comisión de Valores y Bolsa (SEC, por sus siglas en inglés), alertaba de que la compañía estaba «promoviendo de forma activa (aunque no necesariamente consciente) esa clase de actividades».⁷²

Sin embargo, durante mi visita en Facebook, cualquier pregunta sobre las consecuencias de dirigir una proporción cada vez mayor de la experiencia humana a través de unos algoritmos y unas interfaces que parecen juegos diseñados sobre todo para «maximizar la participación de los usuarios» no obtenía más que miradas de incompreensión. Ejecutivos que tan solo unos minutos antes habían indagado en materias tan sensibles como el terrorismo o la legislación extranjera parpadeaban perplejos y cambiaban de tema como si no hubiesen entendido mis palabras. «La gente utiliza los móviles para hacer maldades. De hecho, quizás todo lo malo que circula por nuestros servicios está sucediendo a través de un móvil —dijo un ejecutivo que trabajaba desde Londres—. ¿Cree usted que no deberíamos utilizar los móviles o que deberíamos ralentizar su adopción porque se estén usando con fines perjudiciales?»

Una parte de mí tenía la esperanza de que me estuviesen mareando aposta, quería pensar que me entendían pero no podían admitir abiertamente los efectos distorsionadores de la realidad y el comportamiento que provocaban sus servicios. Me sentía como si estuviera en la tercera entrega de *2001: Una odisea del espacio*, cuando dos astronautas supervivientes traman qué hacer con HAL, la inteligencia artificial que, después de supervisar su largo viaje, se ha vuelto loca y ha asesinado a sus compañeros de tripulación. Intenté imaginar la escena clave: cuando un astronauta le pregunta al otro cómo podrían detener la matanza de HAL, ese respondería: «¿La matanza de quién?».

Como prueba de las buenas intenciones de la compañía, casi todos los empleados con los que hablé mencionaron la respuesta de Facebook al genocidio de Birmania. Aunque admitían que se había reaccionado tarde, la plataforma había prohibido la infame organización extremista, encabezada por el monje racista Wirathu, que había utilizado Facebook como arma para incitar actos de violencia masiva contra minorías desde 2014. Eso era, sin lugar a dudas, una demostración de madurez. Pero varias organizaciones de supervisión digital de Birmania llevaban meses diciéndome que la propaganda racista de los grupos de odio seguía siendo generalizada en internet. La explicación resultó estar en los archivos de Facebook. Un manual sobre Birmania, formateado como un desenfadado PowerPoint de treinta y dos diapositivas, afirmaba lo contrario de lo

que los ejecutivos de Facebook creían que era su política:⁷³ no debía eliminarse el contenido publicado por el grupo de odio o que lo apoyase, decía el texto. Dedicaba una página entera a ese punto. Era como intervenir en Ruanda y decir por error que las emisoras de radio favorables al genocidio mantuviesen las emisiones. Incluso después de todas las deficiencias de la plataforma, un error tan flagrante, sobre una materia tan delicada, seguía siendo desconcertante. Cuando les comenté eso a los representantes de Facebook, la compañía dijo que corregirían los documentos.

Los debates que había en Silicon Valley sobre cómo emplear su poder —ceder más o menos a los Gobiernos, recalcar la neutralidad o el bienestar social, la coherencia o la flexibilidad— casi nunca consideraban la posibilidad de que no deberían tener tanto poder. Que, en el fondo, unir la información y las relaciones sociales bajo el control de unas compañías que aspiran a la maximización de los beneficios estaba reñido con el bien común.

No obstante, puesto que las críticas a las grandes tecnológicas y las pruebas de sus perjuicios eran tan abundantes que no podían negarse por completo, Silicon Valley optó por un relato interno, durante 2018 y 2019, que permitió a sus directivos sentir que seguían siendo los buenos de la película. Lo llamaron «tiempo bien empleado»,⁷⁴ una expresión prestada de Tristan Harris, el exingeniero de Google que había advertido del condicionamiento adictivo y que dejó la empresa en 2015. Ahora esa era la nueva idea de moda en Silicon Valley. Facebook, Google, Apple y otras empresas introdujeron nuevas funciones para monitorizar y gestionar el tiempo de pantalla de los usuarios. Era una especie de renovación de su imagen: hemos aprendido de nuestros pecados (que hemos reducido oportunamente al concepto «demasiado tiempo de pantalla»), hemos experimentado un gran despertar y ahora somos los defensores del bien.

El Persuasive Tech Lab de la Universidad de Stanford, donde académicos e ingenieros colaboraban para desarrollar servicios lo más adictivos posible, se renombró como «Behavior Design Lab» ('Laboratorio de diseño de la conducta'). Su director tuiteó: «Empezaremos a darnos cuenta de que estar encadenados al móvil es un comportamiento que indica un bajo nivel social, algo parecido al hecho de fumar».⁷⁵ Nir Eyal, el consultor que había sido un precursor a la hora de proponer que las máquinas tragaperras fueran los modelos

de las redes sociales, pasó de ser un gurú de la maximización del tiempo de pantalla a un gurú de la reducción del tiempo de pantalla, y publicó un libro titulado *Indistractable* [Indistraíble].⁷⁶

Harris afirmó que se trataba de una campaña de apropiación que no ayudaba mucho a tratar de resolver los daños reales.⁷⁷ Las aplicaciones siguieron actualizando sus algoritmos pensados para la adicción, pero añadieron pequeños contadores que te decían cuántas horas habías pasado conectado. Otros dijeron que era un truco de *marketing*. A mí me parecía un movimiento para la autoabsolución. Por cada campaña publicitaria que decía a los consumidores que ahora Silicon Valley abogaba por el bienestar digital, había muchos más retiros de yoga para consumo interno o grupos de meditación que les decían eso mismo a los habitantes del valle. Los ejecutivos se reunían para purificar sus almas de culpa, felicitarse por haber evolucionado y luego regresaban para ponerse de nuevo a generar crecimiento a destajo. La autoflagelación como autoafirmación — sentirse bien por sentirse mal— se convirtió en una de las especialidades de Silicon Valley.

«Los directores ejecutivos están sufriendo por dentro. Por la noche no pueden dormir»,⁷⁸ dijo en declaraciones al *New York Times* Ben Tauber, exdirector de producto de Google que había convertido una comuna *hippie* junto al mar llamada Esalen en un centro de retiro para ejecutivos tecnológicos. Se trataba de algo extrañamente rebuscado. Pero era para los ejecutivos lo que la actitud bélica de los directores ejecutivos suponía para la moral empresarial y lo que los moderadores eran para los discursos de odio: una cortina de humo para tapar la brecha no resuelta —y quizás irresoluble— entre el propósito de libertad y revolución declarado por las plataformas y sus efectos reales sobre el mundo.

Me di cuenta de que ahí residía el verdadero problema de gobernanza. Si era un tabú pensar que las propias redes sociales — como en el caso de los cigarrillos— podían ser los causantes de los daños que parecían surgir de forma sistemática al adoptarlas, entonces los empleados encargados de gestionar esos daños estaban muy limitados. Eso explicaba muy bien la extraña incoherencia de los reglamentos. Sin entender del todo el impacto de las plataformas, la mayoría de las políticas eran respuestas reactivas a crisis o problemas: un rumor viral, una oleada de insultos, unos disturbios. Los empleados

con mayor responsabilidad hacían una modificación, esperaban a ver qué sucedía y luego volvían a aplicar modificaciones, como si estuvieran reparando un avión en pleno vuelo.

Para los moderadores que tenían que ejecutar esos planes, la situación no cambió mucho. Siguieron apareciendo noticias de unas condiciones laborales asfixiantes. Jacob, como ocurre a veces con esas fuentes, dejó de comunicarse con nosotros cuando su empresa empezó a buscar al topo que la había avergonzado ante Facebook. En 2018, un moderador estadounidense presentó una demanda contra Facebook,⁷⁹ a la que luego se sumaron varios moderadores más, por no proporcionar las protecciones de seguridad mínimas por ley mientras les obligaba a ver material que la compañía sabía que podía traumatizar. En 2020, Facebook resolvió el caso como una demanda colectiva y aceptó pagar 52 millones de dólares a 11.250 moderadores actuales y pasados de Estados Unidos.⁸⁰ Los moderadores de fuera de Estados Unidos no recibieron nada. El modelo de negocio subyacente no ha cambiado.

La dictadura de los «me gusta»

PRESIDENTE YOUTUBE

Habían transcurrido siete años, recordaba la psicóloga Tatiana Lionço con la voz tensa, desde que un vídeo viral de YouTube le había destrozado la vida. En 2012, Lionço había participado en una mesa redonda sobre cómo combatir la homofobia en las escuelas. Dijo ante un pequeño público de académicos y legisladores que debería convencerse a los padres de que no había nada extraño en que los niños mostrasen curiosidad por los cuerpos y la ropa de sus compañeros.

Poco después, un parlamentario de extrema derecha editó unas imágenes del acto y reordenó las palabras de Lionço para que pareciera que había fomentado la homosexualidad y el sexo entre niños.¹ Aquel diputado, considerado por casi todo el mundo un excéntrico marginal, tenía pocos aliados políticos y poco poder directo. Pero tenía un número sustancial de seguidores en YouTube, donde publicaba fragmentos de vídeo editados. Varios *youtubers* de ultraderecha, entonces una comunidad pequeña pero activa, compartieron el vídeo editado y añadieron sus comentarios, llenos de desinformación. Lionço representaba, decían, una conspiración internacional homosexual-comunista. La psicóloga había apoyado la pedofilia. Distribuía «kits gays» para que los usaran los colegios para convertir a los niños a la homosexualidad. Las afirmaciones de los *youtubers* se extendieron a Twitter y Facebook. Los comentarios a los vídeos se llenaron de llamamientos al asesinato de Lionço.

De entrada, los amigos y compañeros de Lionço no prestaron atención al vídeo por considerarlo ruido salido de las redes sociales. Hasta que la historia inventada se convirtió en una realidad por consenso en las plataformas, lo cual indignó a ciudadanos corrientes. Muchos llamaron a su universidad exigiendo que la echaran. Acusaban a la institución, y a cualquier persona que le mostraba su apoyo, de

poner en peligro a los niños. Sus amigos y compañeros del trabajo se distanciaron de ella.

«Me quedé totalmente sola», me contó Lionço. Hizo una pausa, con el rostro tenso, y bajó la mirada. Quizás las personas de su entorno se sentían avergonzadas por haber permitido que aquello sucediera, me dijo. «Creo que a la gente le da miedo que pueda ocurrirles a ellos.» Incluso tras haberse retirado casi del todo de la vida pública, los *youtubers* de ultraderecha, cuyas audiencias se estaban disparando, siguieron alimentando la historia de la profesora universitaria comunista que planeaba sexualizar a los niños. Aunque al final Lionço volvió a dar clase, su vida nunca ha vuelto a ser la misma, perseguida por la infamia dondequiera que vaya. Las amenazas de muerte son una presencia constante, además de los murmullos de sospecha incluso por parte de compañeros que comparten sus ideas. «Estoy agotada. Han sido siete años —dijo cubriéndose la cara con las manos—. Aquello me rompió. Esa fue la peor parte. Me siento sola.»

Lionço es brasileña. En otoño de 2018, el diputado ultra y *youtuber* que había iniciado la campaña de desinformación contra ella seis años antes —un tal Jair Bolsonaro— se presentó como candidato a la presidencia del país. Todo el mundo preveía que perdería los comicios, pero ganó con una victoria aplastante por diez puntos. Fue el hecho más relevante en política internacional desde la elección de Donald Trump. El sexto país más grande del mundo sería gobernado por un conspiracionista de extrema derecha. Bolsonaro permitió la destrucción de miles y miles de hectáreas de selva amazónica,²mostró su apoyo a la violencia de extrema derecha, atacó de forma implacable las instituciones democráticas brasileñas y destripó su burocracia.

Su ascenso parecía ser consecuencia de una indignación social derivada de la corrupción gubernamental, la crisis económica y el retroceso democrático. Pero los brasileños y los analistas con quienes hablé no dejaban de mencionar las redes sociales estadounidenses. «La derecha brasileña apenas existía hace tan solo dos años —me contó Brian Winter, director de la revista sobre política *Americas Quarterly*—. Salió casi de la nada.» El *establishment* había rechazado a Bolsonaro por sus conspiraciones extremistas, sus discursos de odio y su hostilidad contra las mujeres («No te violaría porque no te lo mereces»,³le dijo una vez a otra diputada). Pero ese comportamiento

pensado para llamar la atención daba buenos resultados en internet. Las redes sociales —y YouTube en concreto—, me dijo Winter, presentaban Bolsonaro como «un personaje reinventado». Antes de las elecciones, Winter había visitado la oficina de Bolsonaro con la esperanza de entender su extraño y repentino ascenso. Los ocho miembros de su equipo «se dedicaron a las redes sociales todo el tiempo que estuve allí —me dijo—. Allí no se hacían labores legislativas».

Y no era solo Bolsonaro. Por motivos que nadie entendía del todo, en Brasil habían proliferado las conspiraciones y nuevas causas radicales que parecían remontarse todas ellas a YouTube. «Empecé a fijarme en YouTube cuando, en el primer debate presidencial, uno de los candidatos habló de la URSAL», dijo la analista Luiza Bandeira, que trabaja en Brasil. La URSAL, un plan ficticio para unificar la América Latina en un superestado pancomunista, se había ido enconando en los márgenes de la extrema derecha brasileña hasta que, a principios de 2016, había explotado en YouTube.⁴ Los vídeos que apoyaban esa afirmación recibían cientos de miles de visualizaciones, motivadas en buena medida —según concluyó Bandeira— por los algoritmos de la plataforma, que de forma sistemática la dirigían incluso a ella de vídeos políticos a vídeos de conspiraciones.

Otras muchas conspiraciones se filtraron hasta el discurso mayoritario, lo que generó ondas expansivas de confusión y miedo por el país.⁵ Un juez que investigaba a políticos de izquierdas que había fallecido en un accidente —se afirmaba en una conspiración— en realidad había sido asesinado. El ejército estaba preparando un golpe de Estado. Agentes extranjeros estaban alterando las elecciones, sembrando enfermedades letales por el país y sobornando al Gobierno para que renunciase a la Copa del Mundo.

Esas historias predominaban en Facebook y Twitter —descubrió Bandeira— y en los meses previos a las elecciones de 2018 tuvieron más visualizaciones que la verdad en ambas plataformas. A menudo procedían de YouTube. Casi todas estaban alineadas con la política de extrema derecha de Bolsonaro y su mentalidad de paranoia y odio, lo que arrastró a los brasileños en esa dirección. «Bolsonaro siempre ha defendido estas opiniones ultrapartidistas, pero nunca habían ganado muchos adeptos —dijo Roberta Braga, compañera de Bandeira en el Laboratorio de Investigación Digital Forense, que monitoriza la

desinformación—. Ahora son algo generalizado.» Las elecciones en las que Bolsonaro accedió a la presidencia también permitieron a dos *youtubers* de extrema derecha entrar en el parlamento federal y a otros muchos a cargos en los estados. A otros se les concederían trabajos en la gestión de las políticas federales. Era como si estuviera formándose una gran ola que empujaba a un subgrupo muy concreto de la derecha radical —los *youtubers*— a las altas esferas del poder.

Brasil, el segundo mercado más grande de YouTube,⁶ representaba un precedente de la influencia creciente de la plataforma, como lo habían sido Birmania, Sri Lanka y Alemania en el caso de Facebook. También proporcionaba un prisma para entender los efectos de las redes sociales en Estados Unidos, que guardan una extraordinaria similitud con Brasil: una democracia presidencial que tiene mucho poder, con una gran clase media, definida por la división racial, una polarización en aumento y una derecha populista creciente que parecía vivir en internet. Y, por encima de todo, daba una idea aproximada de cómo sería el futuro.

Si bien a Trump le habían ayudado las plataformas, no había sido producto de ellas. En Brasil era como si las propias redes sociales hubieran llegado a la presidencia. El país parecía representar un extraño nuevo tipo de orden social y político guiado digitalmente que ya empezaba a vislumbrarse en Estados Unidos al acercarse las elecciones de 2020. Visto en retrospectiva, el Brasil de 2019 anticipaba no solo buena parte del caos estadounidense del año siguiente, sino un futuro para el mundo democrático en general que, si no cambia algo, aún puede hacerse realidad.

Aterricé en el país —colaborando de nuevo con mi compañera Amanda Taub y ahora acompañados de un equipo para rodar un documental— en abril de 2019, tres meses después de que Bolsonaro llegase a la presidencia.

DEMOCRACIA POR ALGORITMO

Cuando Matheus Dominguez tenía dieciséis años, YouTube le recomendó un vídeo que le cambiaría la vida. Había creado un grupo de música con algunos amigos en Niterói, su ciudad natal, situada al otro lado de la bahía de Río de Janeiro. Para practicar, veía tutoriales

de guitarra por internet. Un día, la plataforma lo dirigió a un profesor de música llamado Nando Moura, quien colgaba vídeos caseros sobre heavy metal y, cada vez más, sobre política. Moura, un hombre extravagante y de pelo largo, acusaba a las feministas, los maestros y los políticos de planear el adoctrinamiento de los brasileños en el comunismo y la homosexualidad. Hacía bromas y vestía disfraces estrafalarios. Tocaba *riffs* de guitarra y jugaba a videojuegos.

Dominguez se enganchó al canal. A medida que pasaba más tiempo en la plataforma, YouTube le recomendaba vídeos de otros blogueros de extrema derecha. La mayor parte eran jóvenes, salvo uno: el trajeado Bolsonaro, a quien Dominguez vio por primera vez como invitado en uno de los vídeos de Moura. Eso fue en 2016. En esa época, a Bolsonaro, que hacía tiempo que era diputado en la versión brasileña de la Cámara de Representantes estadounidense, lo evitaban incluso en su propio partido. Pero YouTube, persiguiendo su objetivo de los mil millones de horas de visualización, acababa de introducir su nueva IA de aprendizaje profundo. En Brasil, los *youtubers* de extrema derecha —el partido real de Bolsonaro— vieron cómo se disparaba su exposición.

«Allí empezó todo», me dijo Dominguez, ahora un desgarbado muchacho de dieciocho años con gafas y coleta, que calificó YouTube como el nuevo hogar de la derecha brasileña. El algoritmo de recomendación había «despertado a los brasileños», afirmó. «Promueve el contenido que va a tener más visualizaciones. No les preocupa que sea de izquierdas o de derechas; ellos lo que quieren es dinero», dijo. Ahora la política era la vida de Dominguez. Se había inscrito al nuevo partido de Bolsonaro. Algún día quería presentarse como candidato. Y colgaba largos vídeos en YouTube, que consideraba el centro de la vida nacional.

Dominguez nos condujo a una marcha a favor de Bolsonaro que recorrería el paseo de mar de Niterói pasando por delante de altos edificios lujosos y se ofreció a presentarnos a miembros del partido. Yo tenía mis dudas de que fueran a ser tan directos; unos profesionales no iban a atribuir sus éxitos a una página web extranjera. Pero me equivocaba. Todas las personas con las que hablé insistieron en que las redes sociales estadounidenses los habían llevado hasta allí.

«Ha sido así con todo el mundo —dijo haciéndose oír entre la multitud Mauricio Martins, el vicepresidente del partido en la

localidad—. La mayoría de la gente que hay aquí vinieron a través de YouTube y las redes sociales.» Incluso él, dijo. Unos años antes, el sistema había reproducido de forma automática un vídeo sobre política, un tema por el que él nunca había mostrado interés hasta entonces. Era de Kim Katagiri, un creador de contenido de derechas. «Antes de eso —dijo— yo no tenía formación ideológica o política.» Pero el vídeo captó su atención y vio más. El algoritmo, dijo, le proporcionó «mi educación política».

Brasileños como Martins y Dominguez expresaban algo que iba mucho más allá de cualquier cosa que hubiesen observado investigadores como Jonas Kaiser o Guillaume Chaslot: que YouTube no solo había creado una comunidad virtual marginal o alterado el consumo de ciertos usuarios, sino que había radicalizado a todo el movimiento conservador de su país, y con tal eficacia que había suplantado casi por entero la política de derechas. A unos cuantos cientos de kilómetros de allí, en Belo Horizonte, una ciudad con una magnífica arquitectura colonial, un equipo de investigadores estaba encorvado delante de sus ordenadores intentando entender si eso era verdad. «En YouTube hay un enorme predominio de los canales de derechas —dijo Virgilio Almeida, el informático de barba canosa que dirigía el laboratorio, en la Universidad Federal de Minas Gerais—. Atraen más usuarios, más comentarios y más interacciones que los canales de izquierdas.» Pero una crisis derivada de la corrupción acababa de hacer caer al Gobierno de izquierdas de Brasil, lo que había trastocado la política. Quizás YouTube solo reflejaba el sentir de los brasileños.

Almeida tenía algunas ideas para aislar la influencia de la plataforma y averiguarlo.⁷ Su equipo se percató de que podían medir la tendencia ideológica de vídeos de política de YouTube analizando los subtítulos. Utilizaron un programa especial para monitorizar tendencias en cuanto al tono y la inclinación política de los vídeos. E hicieron lo mismo con los comentarios que hay debajo de esos vídeos. Descubrieron que, más o menos desde que se había actualizado el algoritmo de YouTube en 2016, los canales de derechas habían visto aumentar su público a un ritmo considerablemente mayor que los otros, con lo cual dominaban el contenido político de la plataforma. Las menciones positivas sobre Bolsonaro se dispararon. Y también las menciones a teorías conspirativas que Bolsonaro difundía. Como había

descubierto Chaslot durante las elecciones francesas de 2017, YouTube tuvo una fuerte inclinación pro-Bolsonaro y pro derecha dura durante un periodo en que los resultados de Bolsonaro en las encuestas permanecían estáticos y bajos. La plataforma no reflejaba tendencias del mundo real. Creaba tendencias propias.

Pero ahora Almeida hizo un nuevo descubrimiento, que hacía pensar —de una forma que otros estudios solo habían podido insinuar— que YouTube no solo fomentaba el interés de los brasileños por ver vídeos de tendencia bolsonarista, sino que en realidad estaba cambiando la política subyacente de los usuarios. El giro a la derecha de la plataforma había comenzado con visualizaciones de vídeos. Los canales que elogiaban a Bolsonaro o que usaban palabras clave de la extrema derecha veían cómo subía su contador de visualizaciones, presumiblemente gracias al algoritmo. Después de eso, los comentarios también viraron hacia la derecha, lo cual hacía pensar que YouTube estaba arrastrando a los usuarios hacia el bolsonarismo y no al revés.

Sin embargo, la influencia de las plataformas sobre el terreno resultó extenderse más allá de simples cuestiones de preferencia política. A principios de 2017, justo cuando Matheus Dominguez experimentaba su despertar político en internet, un joven y ambicioso concejal de la ciudad llamado Carly Jordy tuvo una idea. Jordy, un conservador radical en una población de tendencia izquierdista como Niterói, tenía pocas vías obvias para salir de la oscuridad. De carácter rudo y con muchos tatuajes —en la mano izquierda lleva una calavera flameante con ojos de diamante—, se sentía despreciado por la clase política brasileña. Pero en YouTube encontró la compañía de personas que pensaban igual. Veía a Nando Moura, a Jair Bolsonaro y a Jordan Peterson, el psicólogo que se había convertido en una puerta de entrada algorítmica a la derecha alternativa.

El plan de Jordy tomó la inspiración de *youtubers* de extrema derecha de Brasil que habían empezado a animar a los estudiantes a que grabasen a escondidas a sus maestros para demostrar el adoctrinamiento comunista-homosexual,⁸ una conspiración fundamentada en la misma acusación ficticia que Bolsonaro había hecho contra Tatiana Lionço, que habían ido urdiendo en sus canales durante meses. Tenía un número discreto de seguidores en YouTube, donde colgaba vídeos en los que se le veía haciendo callar a gritos a

políticos de izquierdas en su ayuntamiento. Al igual que sus héroes de YouTube, animaba a los escolares de Niterói a grabar a sus maestros. Algunos le mandaban grabaciones temblorosas de sus clases hechas con el móvil, que él editaba y subía a YouTube.

Durante una entrevista en su despacho, Jordy nos enseñó con orgullo la más viral de esas grabaciones. En ella, una estudiante interrumpía a su profesora de historia durante una clase sobre la Alemania nazi. La estudiante preguntaba si sus compañeros de clase conservadores eran parecidos a los nazis, a lo cual la profesora respondía que sí. En una imagen sobreimpresionada aparecía el nombre y los usuarios en redes sociales de Jordy. Resultó que había editado las imágenes para engañar. En realidad, la estudiante, que era lesbiana, había contado que la acosaban algunos compañeros de clase cuyas familias apoyaban a Bolsonaro, quien aún era un personaje local conocido por haber dicho que preferiría que su hijo estuviera muerto a que fuera gay. La estudiante había preguntado si esa opinión hacía que sus compañeros de clase fueron como los nazis. No, había dicho la profesora, aunque ambos grupos compartían la homofobia.

Al principio el vídeo tuvo poco público; hasta que lo compartió Nando Moura, quien dijo que era una prueba de que ese colegio era un epicentro de abusos. La grabación se propagó enseguida por los canales de derecha de la plataforma y llegó a una enorme cantidad de usuarios. Luego saltó a Facebook, donde obtuvo cinco millones de visualizaciones.

Valeria Borges, la profesora que aparecía en el vídeo, describió los meses posteriores a la difusión del vídeo como «el peor periodo de mi vida». Borges nos enseñó páginas de amenazas gráficas que le habían llegado. Otros profesores y trabajadores del colegio también habían recibido ataques al ser considerados presuntos cómplices. Aunque los estudiantes de Borges le mostraron su apoyo, la mayoría de los brasileños solo la veían como una persona malvada. «Me ven como a un enemigo, como alguien que debe ser destruido», dijo. Y añadió que, dos años después, aún «tengo miedo por mi familia, por mis estudiantes y por mis compañeros».

Otro profesor del colegio nos contó que los padres ya no estaban seguros de qué creer. Sabían que el colegio y los profesores eran competentes y simpáticos. Pero eran tantos los amigos y parientes que les habían repetido la conspiración, que les costaba conciliar la

realidad que conocían con la irrealidad de las redes sociales, que la había superado. Algunos empezaron a hacer preguntas incisivas sobre el «adoctrinamiento». Los profesores dijeron que tenían más miedo y cautela en sus aulas. Les preocupaba que cualquier cosa que dijese pudiera terminar en internet, distorsionada para favorecer algún programa político, potenciado en todo el país por las redes sociales.

Se estaba propagando por Brasil una oleada de incidentes de esa índole.⁹ La acusación inicial era estrafularia —un profesor había animado a sus alumnos a traficar con drogas, a iniciar una insurrección comunista o a «volverse» gais—, pero supuestamente probada con un vídeo. Otros *youtubers* aportaban su toque al vídeo; así obtenían cientos de miles de visualizaciones (y un porcentaje de los ingresos publicitarios) y suministraban a la plataforma las conspiraciones basadas en guerras culturales que YouTube recompensaba de forma sistemática. Cada rumor aislado, por improbable que fuera por sí solo, daba credibilidad a los demás, una celosía autosostenida de desinformación. Algunos profesores eran despedidos o tenían que esconderse por el acoso, se asediaban colegios y se inundaba a comunidades de desconfianza y división, todo ello orquestado en YouTube.

«En Brasil, esto ahora sucede a menudo, lo de las grabaciones y los linchamientos —dijo Borges—. Es una forma de intimidación. Y funciona.»

Jordy publicó vídeos de otros profesores, todos editados de la misma forma. Le permitían llegar, según él, a un «público nacional». Después de tan solo dos años en el ayuntamiento, en las mismas elecciones que encumbraron a Bolsonaro, obtuvo un escaño en la Cámara de Diputados. Y otro *youtuber* de extrema derecha que había difundido acusaciones de adoctrinamiento en las escuelas fue elegido para el Senado. Era, visto en retrospectiva, una señal de alarma. Los conspiracionistas estadounidenses marginales con aspiraciones políticas, como Lauren Boebert y Marjorie Taylor Green, aunque parecieran ridículos, en realidad podían representar una tendencia creciente cuya culminación podría tener lugar dentro de unos años. «Si no existiesen las redes sociales, yo no estaría aquí —me dijo Jordy—. Y Jair Bolsonaro no sería presidente.»

Cuando le pregunté a Jordy por sus ediciones del vídeo, con las que había alterado su significado, él no me rebatió. Le leí algunas de

las amenazas que le habían llegado a Borges: promesas de que iban a matarla o peor. ¿Se arrepentía de trastocar la vida de unos profesores que sabía que eran inocentes?

El equipo de rodaje había pedido a los trabajadores de Jordy que apagasen los aires acondicionados porque causaban interferencias con el audio. Fuera hacía casi treinta y ocho grados y había mucha humedad. Estábamos cinco personas en su diminuto despacho, casi tocándonos. Llevábamos más de una hora hablando, y el ambiente, viciado y caluroso, era sofocante.

Jordy, con el sudor cayéndole por la frente, habló a las claras. «Lo hice para escandalizar, lo hice para desenmascararla —dijo levantando los hombros e hinchando el pecho—. Quería que pasase miedo.»

A lo largo de todo el incidente, le dije, no había hablado nunca con estudiantes o asociaciones de padres en cuyo nombre afirmaba actuar. Nunca visitó el colegio, que estaba a pocos metros de su despacho. Nunca hizo uso siquiera de sus competencias como concejal. Había centrado todas sus energías en las redes sociales. ¿Por qué?

«Estamos librando una guerra cultural —aseguró, aunque no concretó contra quién—. Las personas respetan a lo que temen. Deben tener ese miedo para entender que se les puede castigar por sus actos.»

Parecía no tener muchos puntos programáticos más allá de atizar la indignación y captar la atención en las redes sociales, cuyos códigos e incentivos, a fin de cuentas, había seguido para llegar a un cargo importante. A medida que hablaba con habitantes de la ciudad, menos parecía que las personas como Jordy y Bolsonaro dirigiesen las fuerzas digitales que los habían encumbrado; más bien parecían beneficiarios pasivos.

«Aquí un 95 % de los niños utilizan YouTube. Es la fuente principal por la que consumen información», me contó Inzaghi, un estudiante de diecisiete años, fuera del colegio que salía en el vídeo de Jordy. Veían YouTube para pasar el rato en el autobús, me dijo, en lugar de ver la tele en casa, incluso para buscar información al hacer los deberes. Pero él y sus amigos se habían dado cuenta de lo mismo que el resto de los ciudadanos de su país. «A veces estoy viendo vídeos sobre un juego y de repente sale un vídeo de Bolsonaro —dijo—. [El sistema] va a intentar que veas esos vídeos tanto si quieres o como si

no.»

Él se dio cuenta de su influencia también en clase. Cada vez más, los alumnos interrumpían a los profesores con acusaciones o conspiraciones que habían oído en YouTube. «Cada vez que alguien dice algo extremo es porque citan a personas como Mamãefalei, Kim Kataguiiri y MBL», dijo otro estudiante, Jojo, listando *youtubers* de derechas. «Nando Moura, Nando Moura —gritó Inzaghi, mientras sus amigos se reían—. En la calle, en el autobús, en grandes grupos veo a gente viendo vídeos de Nando Moura.» Todos los estudiantes con los que hablé destacaron que les gustaba YouTube, sobre todo los canales de juegos y humor. Pero todos se quejaban de que la plataforma los empujaba, una y otra vez, a ver conspiraciones y diatribas políticas. «Termina afectando a la forma de pensar de la gente», dijo Inzaghi.

Pedí a Jonas Kaiser y Adrian Rauchfleisch que me ayudasen a entender el comportamiento de YouTube en Brasil. Ellos repitieron sus métodos perfeccionados en Alemania y Estados Unidos. Rastreando muchas veces las recomendaciones del sistema, descubrieron, como en los otros casos, una agrupación enorme de canales de extrema derecha y conspiraciones y generada por el algoritmo. Una vez más, el sistema utilizaba voces moderadas como puntos de entrada para arrastrar a los usuarios hacia los extremistas y luego los mantenía allí e iba mostrándoles más y más vídeos. Si quedaba alguna duda sobre si Alemania o Estados Unidos podían suponer una excepción, Brasil la despejó: los métodos de YouTube que favorecían la radicalización de derechas a gran escala eran tan estremecedoramente sistemáticos que tenían que estar codificados en la plataforma.

Las consecuencias del efecto madriguera de YouTube en el mundo real —la distorsión impuesta a millones de ciudadanos, una parte sustancial del electorado— estaban por todas partes, empezando por la política. Bolsonaro animaba a los ciudadanos a ver YouTube en lugar de ver noticias respetables.¹⁰ Sustituía a tecnócratas del Gobierno por personajes de las redes sociales,¹¹ quienes utilizaban su poder para reaccionar ante las conspiraciones estrafularias —sobre educación, salud pública o minorías— que habían amansado al algoritmo de Silicon Valley que los había llevado hasta allí. Y los *youtubers* que siguieron a Bolsonaro hasta posiciones de poder continuaron publicando frenéticamente en la plataforma, pues entendían que dependían de ella para mantener a los votantes de las redes sociales

que constituían su base electoral. Eso significaba complacer, en todo lo que hacían desde sus cargos, las necesidades y los sesgos de la plataforma.

La zona cero de esa nueva era de la política hecha por YouTube era la sede central en São Paulo del Movimiento Brasil Libre (MBL). Kim Kataguiiri, máximo responsable de la organización, había alcanzado la fama en YouTube de adolescente publicando vídeos en los que refutaba lo que consideraba los sesgos izquierdistas de sus profesores. (Maurício Martins, el miembro del partido de Bolsonaro, había citado uno de esos vídeos —que YouTube le había reproducido de forma automática— como su introducción a la política.) En 2016, él y otras personas fundaron el MBL para hacer campaña a favor del proceso de destitución política de la entonces presidenta de Brasil. Los miembros de la organización eran en su mayoría jóvenes, formados, de derechas y muy presentes en la red: el grupo demográfico de YouTube. Ese otoño, a los veintidós años, Kataguiiri obtuvo un escaño en la Cámara de Diputados.

Ahora, tras algunos meses de legislatura, Kataguiiri se pasó por las oficinas de la organización, un garito de suelos de hormigón con hípsters encorvados delante de portátiles. Un cámara con una gorra de Make America Great Again puesta del revés le indicó que fuera a un sofá de piel, donde grabó el vídeo del día. Después de filmarlo, Kataguiiri me contó que al principio el MBL usaba Facebook, pero que después de 2016, más o menos cuando YouTube actualizó su algoritmo, la plataforma de vídeos había demostrado ser más efectiva, y entonces cambiaron. Además, añadió, Facebook había tomado medidas varias veces contra su organización por desinformación,¹² mientras que YouTube nunca lo hizo, a pesar de que el MBL colgaba el mismo contenido en ambas plataformas. Eso era algo que oía reiteradamente: YouTube era mucho más permisivo, uno de los motivos por los que les encantaba a organizaciones como la suya.

Otros *youtubers* del MBL también obtuvieron representación en 2018. Sus canales ocupaban todos el nodo central del mapa de Kaiser sobre el YouTube brasileño, potenciado con gran fuerza por las recomendaciones del sitio. Uno de ellos, Arthur do Val, incluso tras obtener un escaño en el parlamento de su estado, seguía utilizando el nombre de su canal de YouTube, «Mamãefalei» ('Mamá dijo').

Para subrayar la importancia de la plataforma, Do Val citó un vídeo superviral que otra organización acababa de publicar. Titulado «1964», en referencia al año del golpe de Estado militar de Brasil, afirmaba que los abusos de la dictadura —que había asesinado a cientos de disidentes y torturado a miles— se los habían inventado los historiadores de izquierdas. Sostenía que el golpe había sido necesario para acabar con el comunismo, con lo que se daba a entender que pronto podría requerirse una acción como aquella. «Yo fui una de las personas que más compartió ese vídeo», dijo Do Val.

El vídeo «1964» aparecía por todos los lugares que visitaba de Brasil. Dominguez, el activista de dieciocho años, lo calificaba como «un hito» que lo había convencido de que el régimen militar no había sido tan malo. Valeria Borges, la profesora de historia, dijo que era «espantoso» ver que YouTube borraba la historia de su país de las mentes de los jóvenes brasileños, o incluso que los impulsaba a repetirla.

Mientras estábamos en la sede central del MBL, los activistas admitieron que, pese a su éxito, incluso ellos habían llegado a preocuparse por los efectos de las plataformas. «Aquí tenemos algo que llamamos la dictadura de los “me gusta”», dijo Pedro D'Eyrot, un exguitarrista de rock con moño que se había politizado y ahora hacía vídeos para el MBL. Habían visto, nos dijo, cómo un *youtuber* tras otro se volvían gradualmente más extremos, más falsos y más temerarios, «solo porque algo te va a dar visualizaciones, va a generar interacción».

Todo el mundo sentía esta presión, dijo. «Una vez que abres esa puerta, no hay vuelta atrás, porque siempre tienes que ir más allá.» Antes creía que las redes sociales eran una fuerza para el cambio, dijo. Ahora parecían crear tendencias que solo podían ser dañinas, demasiado extremas incluso para una organización de revisionistas del golpe de Estado que llevan una gorra trumpista. «Terraplanistas, antivacunas, teóricos de la conspiración en política: todo es el mismo fenómeno —dijo—. Lo ves en todas partes.» Si algunos de los mayores explotadores y beneficiarios de ese sistema hablaban de repente como Renée DiResta o Guillaume Chaslot, entonces los peligros, pensé, debían de haberse extendido más allá incluso de lo que había visto hasta ese momento, aquí o en Estados Unidos. Y así era.

En Maceió, una de las ciudades más pobres de Brasil, en un patio de hormigón en las afueras de la localidad, quince madres se apiñaban debajo de un toldo de plástico resguardándose de la lluvia, esperando a Mardjane Nunes. Desde 2015, miles de mujeres embarazadas de América, infectadas por un nuevo virus llamado zika, han dado a luz a hijos con discapacidades neurológicas graves y cráneos deformes, una enfermedad conocida como microcefalia. El nordeste de Brasil, donde se encuentra Maceió, fue de las zonas con mayor afectación. En barrios como ese, las madres crearon grupos de apoyo para ayudarse unas a otras a gestionar esa enfermedad —de la que se sabía poco— que había dejado a sus hijos discapacitados.

Nunes, destacada experta en el zika, había venido para responder preguntas. En pocos minutos, una madre se levantó, con su hijo en brazos, y dijo que había oído en las redes sociales que el zika no lo causaba un virus transmitido por los mosquitos, como decían los médicos, sino la administración de vacunas caducadas. Atormentada por la incertidumbre, se planteaba no vacunar más a sus hijos. Muchas madres del grupo asintieron con la cabeza. Nunes le dijo, con tono amable pero con total claridad, que esos rumores eran falsos. No obstante, una tras otra, las madres expresaron sus dudas. Todas habían visto eso, o algo parecido, en internet. El zika era una mentira o un complot. No se podía confiar en los médicos. Las vacunas no eran seguras.

A Nunes ya le habían contado compañeros suyos que esas reuniones solían ir así, dijo después. Poco antes, había dejado un puesto de trabajo respetado en el Ministerio de Sanidad para volver a las tareas en primera línea, combatiendo el VIH y el zika en el hospital local y mediante grupos comunitarios. Las redes sociales, dijo, se estaban volviendo un motivo de alarma creciente en el ámbito de la salud pública. Los investigadores sobre el terreno recogían historias de poblaciones enteras que rechazaban tratamientos, madres aterrorizadas hasta el punto de negar a sus hijos cuidados que les salvarían la vida. Pero ella no estaba preparada para ver aquel miedo en los ojos de las madres, toda una comunidad escindida de la realidad. «Si este grupo, que ya está más involucrado, en el que hay comunicación entre sus miembros, aún tiene dudas —dijo—,

imaginaos qué les pasará por la cabeza a las madres que no cuentan con un grupo como este.»

Eso explicaba, añadió, por qué su ministerio había destinado tantos recursos a educar a familias jóvenes como esas, aunque lo había terminado superando la desinformación que se propagaba aún más deprisa por las redes sociales. El rechazo a las vacunas iba en aumento en Brasil.¹³ Como también las noticias de comunidades que se negaban a usar larvicidas antimosquitos, uno de los métodos preferidos para combatir el zika que a menudo se menciona en vídeos conspirativos.¹⁴ «Las redes sociales están ganando», dijo Nunes.

Ese problema no era nuevo. Ya en 2015 DiResta había descubierto que Facebook fomentaba conspiraciones sobre el zika de forma sistemática. El Instituto Brasileño de Investigación y Análisis de Datos descubrió que, en 2017, YouTube se había llenado de conspiraciones sobre la vacuna de la fiebre amarilla.¹⁵ Los casos de fiebre amarilla aumentaron alrededor de esa época, aunque no estaba claro qué era el causante de qué. Fuera como fuese, muchas de esas conspiraciones habían sido propagadas por canales de extrema derecha. Esa era otra prueba de la conclusión de Kaiser de que el algoritmo de YouTube, al enlazar canales de conspiraciones y extremistas antes desvinculados, los cruzaba para crear una amenaza de nuevo cuño.

«Todo lo que no sabes lo puedes encontrar en YouTube», le dijo al día siguiente Gisleangela Oliveira dos Santos, una de las madres que había hecho preguntas sobre las vacunas, a mi compañera Amanda en su casa. Pese a que Oliveira dos Santos vivía en un lugar que estaba a un día de camino, Amanda, también madre joven, insistió en que el equipo de rodaje hiciera el viaje con ella, con lo cual pudimos ver que los padres que reaccionan a esas conspiraciones, aunque es fácil confundirlos como parte del problema, también son víctimas de las redes sociales.

Cuando al bebé de Oliveira dos Santos le habían diagnosticado microcefalia tres años antes, la información era escasa. El virus, al fin y al cabo, era nuevo. Ella buscó todo lo que pudo en internet, también en YouTube.

Una y otra vez, la plataforma le servía vídeos responsabilizando del zika a las vacunas del sarampión caducadas que el Gobierno había comprado a bajo coste. Era culpa del mercurio que había en las

agujas. De una camarilla de extranjeros turbios que querían debilitar a las familias católicas. En algunos vídeos se había creado un decorado para que pareciese un informativo televisivo o un anuncio de los servicios públicos. En algunos aparecía un sacerdote u otro personaje de confianza implorando a las buenas madres brasileñas que no escuchasen a los médicos, corruptos y controlados desde el extranjero. Las otras madres de su comunidad tuvieron la misma experiencia, y compartían los vídeos en grupos de WhatsApp, otra red de confirmación entre grupos de iguales.

Oliveira dos Santos había terminado viendo esos vídeos, dijo, por el sistema de recomendaciones de YouTube y su motor de búsqueda. Tan solo buscar en Google términos como «zika» o «vacunas contra el zika» la llevaba a esos vídeos, que Google —una vez más con su favoritismo sinérgico de los enlaces de YouTube que habían resultado ser tan lucrativos— a menudo ponía en los primeros puestos de los resultados de búsqueda. Oliveira dos Santos sabía que a veces no se podía confiar en internet. Pero los vídeos la hacían dudar y la paralizaron. Al principio, le había puesto a su hija las inoculaciones infantiles estándar. Pero «después de aquello —confesó— me daba miedo ponerle más vacunas a mi hija». Ella también dejó de confiar en la vacunación.

Desde entonces, la han embargado las dudas y, sobre todo, la culpabilidad. ¿Había provocado ella la enfermedad potencialmente mortal de su hija al aceptar las vacunas? Los médicos decían que no. Pero esas compañías tecnológicas estadounidenses, muy respetadas en Brasil, le decían que sí todo el tiempo. Incluso le mostraban «pruebas» y se las enseñaban casi cada vez que buscaba información sobre cómo cuidar a aquella hija que YouTube y Google insistían en que habían dejado discapacitada de por vida. «Siento impotencia —nos dijo—. Me siento impotente y herida.» (Como otros muchos brasileños, cuanto más tiempo pasaba en YouTube, más vídeos veía que apoyaban a Bolsonaro. Le parecieron convincentes, dijo, y le votó.)

El hospital público de Maceió no cuenta con una dotación de fondos especialmente cuantiosa. Cuando lo visité, muchas de las luces del techo estaban apagadas para ahorrar electricidad. Pero su lugar en la línea de combate contra el zika había atraído algunos de los profesionales médicos con mayor talento del país. Cuando los médicos del centro supieron por qué visitaba yo el sitio, enseguida me hicieron

un hueco en sus agendas y me condujeron a la sala de descanso para el personal, que casi estaba vacía. «Las noticias falsas son una guerra virtual. Nos llegan desde todas las direcciones», dijo el neurólogo pediátrico Flavio Santana. Los debates dentro de la comunidad médica del país, dijo, giraban cada vez más en torno a la frustración por las redes sociales. «Si va usted a otros lugares de Brasil, se va a encontrar con los mismos problemas.»

Auriene Oliveira, especialista en enfermedades infecciosas, asintió con la cabeza. Cada vez más, los pacientes cuestionaban sus consejos y ponían en duda los hechos. «Dicen: “No, lo he buscado en Google, lo he visto en YouTube”», dijo. Entendía el dilema de sus pacientes. Las conspiraciones presentaban un nivel de certeza que la ciencia no podía ofrecer. Los remedios caseros permitían a las madres sentir que retomaban el control de la salud de sus hijos. Y las plataformas eran omnipresentes mientras que la doctora Oliviera no podía serlo, ya que aparecían día tras día en las vidas de unas madres que, como compaginaban varios trabajos o vivían lejos, solo acudían a la consulta una vez al mes. Así pues, no recurrían a las aplicaciones móviles por pereza o ignorancia, sino por necesidad.

No obstante, los vídeos virales ponían cada vez en mayor riesgo la vida de los niños, pues aconsejaban a las madres no solo que rechazasen las vacunas, sino también que tratasen ciertas enfermedades por su cuenta y que no permitiesen que a sus hijos les pusieran sondas nasogástricas en caso de padecer una dolencia que les afectase la deglución. Algunas veces, contaban los médicos, habían amenazado a las madres con ponerse en contacto con los servicios sociales para forzarlas a escuchar. Se sentían fatal por usar este poder contra esas mujeres, pero a veces era la única forma de contrarrestar la capacidad de influencia de las plataformas.

¿Cuánto tiempo llevaba ocurriendo eso?, pregunté yo. Si había coincidido con la llegada del zika, en 2015, entonces probablemente se debía más a la confusión creada por el virus que a cualquier cambio tecnológico. Pero los médicos nos dieron la misma respuesta que el resto de la gente. Había empezado más adelante, justo después que de YouTube actualizara sus algoritmos. «Más de una vez por semana aparecen este tipo de cosas —dijo Oliveira—. Cada vez es más grave.» Médicos e investigadores lamentaban a menudo, dijo la experta, que competían contra «el doctor Google y el doctor YouTube». Y estaban

perdiendo.

Recopilé los vídeos conspirativos que habían citado las madres y los trabajadores sanitarios —muchos tenían decenas o cientos de miles de visualizaciones— y se los mandé al equipo de Kaiser. ¿Esos vídeos —y los canales que los habían subido— habían aparecido en su análisis de la red?, le pregunté. Resultó que no solo aparecían, sino que el sistema los promocionaba. El equipo identificó una enorme red de canales sobre salud y bienestar enlazados por el algoritmo que versaban sobre toda clase de aspectos, desde los informes médicos hasta los cristales curativos.¹⁶ Como sucedía con los vídeos sobre política, el algoritmo utilizaba canales más creíbles o habituales a modo de puertas de entrada para dirigir a los usuarios hacia las peores conspiraciones y la desinformación. Los mismos vídeos que nos habían enseñado padres y médicos angustiados estaban cerca del centro de esa red, la destinación final del algoritmo.

Eso era una prueba contundente de que todas las personas con las que habíamos hablado sobre la desinformación médica en Brasil tenían incluso más razón acerca de YouTube de lo que pensábamos. De que la plataforma explotaba el interés lógico de los usuarios por los asuntos médicos —como hacía con la política— para arrastrarlos por madrigueras en las que de otra forma nunca se habrían metido. Y de que, del mismo modo en que YouTube había aprendido a secuenciar vídeos de política para convertir a usuarios esporádicos en radicales adictos al consumo digital, había terminado mostrando vídeos sobre el zika y sobre vacunas en el orden adecuado para convencer a unas afectuosas madres de que pusieran a sus hijos en peligro adrede.

«En las plataformas siempre habrá contenido que roce los límites. Esto es previsible —dijo Kaiser esforzándose por simpatizar con las compañías tecnológicas—. Lo chocante —agregó— es que los algoritmos de YouTube están ayudando a los usuarios a ir en esas direcciones.»

EL CONDUCTO

Al enterarse de lo que Amanda y yo estábamos investigando, Luciana Brito, una psicóloga clínica de voz suave que habíamos conocido en el centro para madres de Maceió, insistió en que charlásemos. Estaba de

visita en la ciudad: había venido desde su centro de investigación situado en la capital de Brasil para hacer trabajo de campo con familias afectadas por el zika. Brito estaba ocupada todo el día —los padres estaban en apuros— y no pudimos sentarnos a conversar hasta después de medianoche. En el móvil nos enseñó un mensaje de WhatsApp que había recibido del padre de un niño con microcefalia. Era un vídeo en que se afirmaba que el zika lo había propagado la Fundación Rockefeller en un complot para legalizar el aborto en Brasil. El padre exigía que le dijese si eso era verdad.

Eso, dijo Brito, ahora sucedía todo el tiempo. En muchos lugares del mundo, la gente no se puede permitir tener un ordenador o banda ancha, ni siquiera el gasto de datos asociado a reproducir vídeos. WhatsApp proporcionaba una alternativa. Con programas de datos gratuitos que a menudo cubrían los gastos derivados de usar la aplicación, quienes no podían permitirse ver vídeos en YouTube podían ver fragmentos que se habían subido a WhatsApp y luego reenviarlos a amigos y compartirlos en grupos de WhatsApp de grandes dimensiones. En zonas de Brasil donde el analfabetismo es elevado, se cree que este es uno de los métodos principales con los que muchas familias consumen noticias. Los grupos de WhatsApp son su Google, su Facebook y su CNN, todo en uno.

Como quienes copiaban los fragmentos de los vídeos eran usuarios habituales, era más probable que se volviera a subir a WhatsApp y que se hiciera viral allí lo que estuviera de moda en YouTube (o Facebook), al igual que una infección que salta de un huésped a otro. Los vídeos como el que le había mandado a Brito ese contacto —nos dijo la experta— a menudo se extendían por grupos de WhatsApp que se habían creado para compartir información sobre cómo enfrentarse al zika, lo que ponía en su contra los esfuerzos de los ciudadanos para gestionar la salud de su familia. Brito y sus compañeros se unieron a los grupos, donde intentaban desmontar las peores afirmaciones, pero seguían recibiendo preguntas. Y constantemente se filtraba contenido de YouTube y Facebook sin que nadie lo verificase.

Virgilio Almeida, el experto informático de la Universidad Federal de Minas Gerais, había estado estudiando justo este fenómeno.¹⁷ Él y su equipo rastrearon decenas de miles de mensajes de cientos de grupos de WhatsApp brasileños (todos anonimizados, por

supuesto) y luego buscaron tendencias. Descubrieron que los usuarios de WhatsApp subían un vídeo por cada catorce mensajes de texto, un ritmo altísimo. Los usuarios de WhatsApp también adjuntaban más enlaces a YouTube que a cualquier otra página —con una frecuencia diez veces superior que a Facebook—, lo que reforzaba la teoría de que había un conducto que conectaba YouTube y WhatsApp. Encontraron tendencias parecidas en India e Indonesia, lo que hacía pensar que el efecto podía ser universal.

Aquello era otra cosa, dijo Brito. A medida que se propagaban las conspiraciones sobre el zika, los *youtubers* de extrema derecha se habían adueñado de ellas y les habían dado otra vuelta de tuerca. Las organizaciones por los derechos de las mujeres, afirmaban los *youtubers*, habían contribuido a diseñar el virus como una excusa para imponer abortos obligatorios. A menudo los vídeos mencionaban la organización de Brito. Los espectadores, ya aterrorizados por el zika, tomaban nota. «Justo después de publicarse un vídeo, empezamos a recibir amenazas», dijo Brito. La mujer sabía que las afirmaciones proporcionaban alivio psicológico a las familias, pues les permitían proyectar hacia fuera su culpabilidad y su miedo. Era el constructo crisis-solución de J. M. Berger: del mismo modo en que los *youtubers* estadounidenses habían convertido la desposesión de los hombres jóvenes en rabia contra las feministas, ahora los de Brasil estaban explotando el terror de unas familias que se enfrentaban a una enfermedad implacable.

Las amenazas contra Brito y sus compañeros se habían vuelto tan frecuentes que la policía creó un canal especial para que les informasen si ocurría algo que les parecía grave. Lo usaban alrededor de una vez por semana, dijo Brito. Pero lo que más le preocupaba eran las madres que habían interiorizado las conspiraciones. Cada vez había más madres que rechazaban la ayuda de las organizaciones de apoyo, ya que ahora sospechaban que afligían a sus hijos. Y, sin apoyo médico ni emocional, recurrían más aún a YouTube. «Estas mujeres son muy vulnerables —aseguró Brito—. Es muy fácil que caigan en la trampa de creer esas teorías. Así que hay mucha desesperación.»

Había un nombre que había oído una y otra vez en Maceió por boca de médicos, trabajadores sanitarios y las madres arrastradas por las madrigueras de YouTube. Nadie había visto a Debora Diniz durante meses, pero todo el mundo insistía en que hablase con ella.

Diniz, una mujer delgada con el pelo corto y cano que durante años se había desempeñado como abogada pro derechos humanos, tiempo atrás había empezado a trabajar como directora de documentales. Su película sobre el brote de zika la había llevado a Maceió. Conmovida por las madres, los médicos y los trabajadores humanitarios que habían aunado esfuerzos, se quedó en la zona trabajando en representación de la comunidad.

Un día, las amenazas empezaron a inundarle el móvil y el correo electrónico. Al principio, Diniz las ignoró. Como veterana de las guerras culturales en Brasil —en especial por el aborto— estaba acostumbrada. Pero aumentaron hasta límites insólitos. Y las acusaciones eran disparatadas: que si trabajaba para George Soros, que si había participado en la creación del zika. Muchas citaban la misma fuente: Bernardo Küster.

Küster era una creación de YouTube de pies a cabeza. Desde su casa, en la pequeña ciudad de Londrina, durante bastante tiempo este ultraconservador de treinta años había soltado diatribas contra liberales ateos, el papa, periodistas, la «ideología de género» (es decir, los derechos LGBT) y, en particular, las feministas. Recompensado una y otra vez por los sistemas de YouTube, había acumulado 750.000 suscriptores. Una vez, Bolsonaro había apoyado su canal. Y, desde finales de 2017, Küster había puesto a Diniz en el centro de un complot para convencer a las madres reticentes al aborto de interrumpir el embarazo. Como los vídeos generaban visualizaciones, fue haciendo más, en los que se inventaba nuevas afirmaciones que pusieron en contra de Diniz a sus miles de seguidores, furiosos por los pecados de aquella mujer.

Las amenazas que inundaban el móvil de Diniz se volvieron más frecuentes y más gráficas, con largas descripciones de los planes del remitente de violarla y torturarla. A menudo repetían lo que fuera que hubiera dicho Küster en su último vídeo. Algunos citaban detalles de sus rutinas diarias. «Describen la forma en que me van a matar», me contó Diniz cuando nos encontramos en Nueva York, donde ahora vivía, tras haber huido de Brasil. Mientras el equipo de grabación iba trabajando, ella agarró el móvil y me leyó en voz alta: «La voluntad de Dios es que te mate. Luego me voy a suicidar». En sus vídeos Küster había mencionado las amenazas de paso, pero nunca las había apoyado de forma explícita. YouTube se negaba a eliminar los vídeos,

a pesar de las súplicas de Diniz alegando que ponían en peligro su vida.

Después de meses recibiendo amenazas, la universidad en la que Diniz daba clase recibió un aviso. El remitente decía que se dirigía al centro para abatir a la mujer, luego a sus estudiantes y añadía que a continuación se suicidaría. La policía le dijo a Diniz que ya no podían garantizar su seguridad. A ella le preocupaban también sus compañeros del trabajo y sus padres ancianos, que habían empezado a recibir amenazas parecidas. Entonces se marchó de Brasil sin saber cuándo podría regresar, si es que podía volver algún día.

«La desenmascaré», dijo Küster en un vídeo con muchas visualizaciones en el que se alegraba por su exilio. «Aquí no eres bienvenida», gritaba mientras movía un dedo con actitud amenazadora ante la cámara. Pero Diniz creía que, en muchos sentidos, Küster era producto de unas fuerzas superiores a él. Tanto él como las amenazas que había inspirado —sostenía Diniz— procedían de un «ecosistema de odio» cultivado por YouTube. «El sistema de recomendaciones de YouTube, que te lleva de un vídeo al siguiente, alimenta ese ecosistema», dijo hablando despacio mientras elegía cada palabra en inglés, su tercera lengua. «“En el primer vídeo me dicen que ella es una enemiga de Brasil. En el siguiente me dicen que las feministas están cambiando los valores familiares. Y en el siguiente, que reciben dinero del extranjero”. El bucle es lo que lleva a una persona a decir: “Haré lo que hay que hacer”.» El efecto era tan convincente, dijo, porque «parece que la conexión la haga el usuario, pero la hace el sistema».

Los exilios como el de Diniz eran cada vez más habituales. Al fin y al cabo, uno no puede protegerse ante una amenaza que viene de todas partes. Unos meses antes, Marcia Tiburi, una activista de izquierdas que se había presentado como candidata a gobernadora del estado de Río de Janeiro, huyó de Brasil para escapar de las amenazas de muerte instigadas en gran parte en las redes sociales. Jean Wyllys, el primer diputado abiertamente gay del país, había hecho lo mismo.

La noche antes de la entrevista, nos contó Diniz, había cenado con Wyllys, que estaba en la ciudad. Alguien debió de reconocerlos, porque un grupo de brasileños se congregó fuera del restaurante y empezó a sacar fotos y a señalarlos. La multitud los amenazó mientras se iban y los siguió calle abajo profiriendo insultos y acusaciones que

repetían las conspiraciones de YouTube. Por eso Diniz rara vez salía, ni siquiera allí.

«Estamos ante una milicia impulsada por los algoritmos —dijo—. Los algoritmos están construyendo a la milicia.»

Le pregunté si aún recibía mensajes amenazadores.

«A diario —contestó—. Nunca te sientes cómodo en una situación como esta.»

Llevábamos una hora hablando, las cámaras y las luces estaban al acecho, cuando le pedí que repasara, para una grabación y en una lengua extranjera, su propio trauma. Le dije que no era capaz de imaginar lo que le supondría seguir levantando la voz, sabiendo que los peligros la habían seguido incluso hasta allí.

Ella abrió los brazos encima de la mesa en la que estábamos hablando y bajó la mirada. Estaba llorando. «Hay momentos en tu vida en los que te encuentras ante un dilema —dijo—. Solo tienes dos opciones. Una es dejarles ganar. Y la otra es contraatacar. Yo no quiero que mi vida se rija por este lenguaje. Pero dedicaré mi vida a cambiar esta situación. Y no me van a matar. No me matarán.»

Le pregunté si quería dejar ahí la entrevista, pero ella hizo que no con la cabeza. Dijo que había seguido las historias de médicos y activistas que habían sido víctimas de vídeos de YouTube, de maestros que habían tenido que esconderse, de vidas arruinadas y de comunidades puestas patas arriba. «Tenemos que conseguir que las compañías asuman su responsabilidad», dijo. Instó a los ejecutivos de YouTube a reflexionar sobre su propia involucreción. «Mi esperanza es que entiendan que forman parte de esta comunidad de odio —añadió—. Son responsables desde un punto de vista ético.»

Aun así, YouTube siguió negándose a asumir responsabilidades. Los representantes de la empresa repetían el guion de siempre. Que la compañía se tomaba en serio la seguridad y el bienestar. Pero rechazaban los estudios de la Universidad Harvard y la Universidad Federal de Minas Gerais alegando que no creían que empleasen una metodología adecuada. Sobre estas cuestiones, la compañía decía que había llevado a cabo sus propias investigaciones internas, que habían demostrado que, en una inmensa mayoría, la plataforma promovía contenido veraz y beneficioso. Los representantes, que exigían revisar cientos de páginas de pruebas detrás de los estudios que citábamos Amanda y yo, no proporcionaban datos ni metodologías que

sustentasen sus afirmaciones.

Tras admitir que algunos de los vídeos de desinformación médica que Kaiser había identificado infringían las normas de la plataforma, la compañía los eliminó. Los detalles de la investigación de Kaiser, que yo había enviado a YouTube, implicaban que había cientos de vídeos —tal vez miles— que instaban a las madres a privar a sus hijos de unos tratamientos que les salvarían la vida. Les habíamos mandado solo algunos enlaces a modo de ejemplo. YouTube nunca pidió los demás. La compañía tampoco pidió detalles sobre los vídeos que habían inspirado amenazas de muerte creíbles contra Diniz y otras personas, muchos de los cuales seguían disponibles en su web.

Ese mismo año, YouTube anunció que había hecho cambios en su algoritmo destinados a reducir «la propagación de contenido dudoso e información perjudicial».¹⁸ Pero algunos de esos cambios ya se habían introducido cuando hicimos nuestro trabajo periodístico, lo que planteaba dudas sobre su efectividad. La compañía publicitaba un indicador algo sesgado para medir el éxito: «Una disminución de un 70 % del tiempo de visualización de este contenido procedente de recomendaciones a no suscriptores en Estados Unidos».

En mayo de 2019, dos meses después de regresar de Brasil, recibí un mensaje de texto de Kaiser: me pedía que le llamase de inmediato. Descolgó al primer pitido, le temblaba la voz. Me contó que él y los otros investigadores habían ejecutado más versiones de su programa de rastreo de plataformas en canales brasileños. Y habían encontrado algo tan perturbador que no sabían qué hacer. Me di cuenta de que la llamada de Kaiser no era para un artículo. Era una llamada de auxilio.

CAMPOS DE ENTRENAMIENTO

Christiane no se imaginaba nada de lo que sucedería cuando su hija de diez años y una amiga colgaron un vídeo en el que salían chapoteando en una piscina que había en el patio de una casa. «El vídeo es inocente, no es nada del otro mundo», dijo Christiane, que vive en un barrio residencial de Río de Janeiro. Unos días más tarde, su hija le dio una noticia emocionante: el vídeo tenía miles de visualizaciones. Poco después, tenía 400.000. Era una cifra asombrosa e inexplicable para un vídeo breve y banal que se había subido al canal de

Christiane, que por lo general recibía unas decenas de clics. «Volví a ver el vídeo y me asustó el número de visualizaciones», dijo Christiane.

Y tenía motivos para asustarse. Sin decir nada, el algoritmo de YouTube había seleccionado el vídeo de su hija para un programa perturbador y de gran envergadura. De entre todos sus archivos, estaba recopilando decenas de vídeos de niños preadolescentes semidesnudos. Muchos los sacaba de películas caseras de familias desconocedoras de lo que estaba ocurriendo. Los juntaba todos y, uno tras otro, iba mostrando vídeos de niños de seis o siete años en bañador o en ropa interior, haciendo el espagat o tumbados en la cama, para atraer a una clase muy específica de espectador con un contenido que encontrarían irresistible. Y luego creaba un público para esos vídeos del tamaño de diez estadios de fútbol. «Estoy estupefacta», dijo Christiane cuando se enteró de lo que había sucedido, aterrorizada por el hecho de que el vídeo de su hija se hubiera presentado junto a muchos otros, lo que dejaba perturbadoramente claras las intenciones de la plataforma.

Kaiser, junto con Rauchfleisch y Córdova, se había topado con esto mientras trabajaba en un estudio sobre Brasil.¹⁹ A medida que su máquina de prueba iba siguiendo las recomendaciones de YouTube sobre vídeos de temática sexual, el sistema conducía a contenido sexual menos convencional y más extremo. Eso no era chocante de por sí; habían visto el efecto madriguera muchas veces en otros tipos de contenido. Pero algunas de las cadenas de recomendación seguían una progresión inequívoca: cada vídeo conducía a otro donde la mujer que era el personaje central ponía mayor énfasis en la juventud y adoptaba una actitud más erótica. Los vídeos de mujeres hablando de sexo, por ejemplo, llevaban a vídeos de mujeres en ropa interior o dando de mamar, y algunas veces se mencionaba su edad: diecinueve, dieciocho o incluso dieciséis. Algunas buscaban *sugar daddies*, un término para referirse a los espectadores excitados que hacen donativos. Otras insinuaban que disponían de vídeos privados donde posaban desnudas a cambio de dinero. Tras algunos clics, la mujer que aparecía en los vídeos adoptaba una actitud prepúber de una forma cada vez más manifiesta hablando como si fuera una niña o posando seductora con ropa infantil.

Desde allí, de repente YouTube pasaba a recomendar vídeos de

niños muy pequeños grabados en momentos de desnudez fortuita. Una niña de, quizás, apenas cinco o seis años cambiándose o contorsionándose para hacer una postura de gimnasia. Y luego una retahíla casi infinita de vídeos de este tipo, procedentes de todo el planeta. No todos parecían ser vídeos caseros; algunos los habían subido cuentas anonimizadas por precaución.

La implacable especificidad de las selecciones de YouTube era casi tan perturbadora como el propio contenido, lo que hacía pensar que sus sistemas podían identificar correctamente un vídeo de un niño semidesnudo y determinar que esa característica era el atractivo de la grabación. Mostrar una serie de vídeos como esos justo después de material sexualmente explícito dejaba claro que el algoritmo trataba a los niños como contenido sexual. El extraordinario número de visualizaciones, a veces de millones, indicaba que eso no era una peculiaridad para unos pocos usuarios. El sistema había encontrado — o quizás creado — un público para los vídeos. Y estaba trabajando para mantener la actividad de ese público.

«Lo que conecta esos canales es el algoritmo de YouTube. Eso es lo que asusta», dijo Kaiser, que calificó las recomendaciones de «perturbadoramente relevantes». Sabía que tenía que actuar con prudencia con esos vídeos. La legislación estadounidense contra el consumo de pornografía infantil deja pocas excepciones a investigadores o periodistas. Probablemente, la mayoría de los vídeos no eran pornografía infantil desde una perspectiva jurídica. Pero algunos sí. Y el contexto en que YouTube los había situado no dejaba espacio para la ambigüedad en cuanto a su contenido. Kaiser y yo acordamos los procedimientos para rastrear los vídeos con responsabilidad y sin exponer a nadie a quien pudiésemos perjudicar. Pero también nos sentíamos obligados a actuar deprisa. Las visualizaciones estaban subiendo a miles por día. Algunos vídeos, como el de la hija de Christiane, estaban subidos con el nombre de los padres, con lo cual los niños que aparecían podían ser localizados con facilidad por algún monstruo lo bastante atrevido. Había que hacer algo.

Juliana Cunha, psicóloga de SaferNet, un organismo brasileño que supervisa internet, dijo que su organización ya había visto antes este fenómeno. Pero nunca con niñas tan pequeñas. Y nunca con un público tan numeroso. Los vídeos de este tipo, comercializados en

foros de la *dark web* y en portales de compartición de archivos, por lo general podían llegar a unos cientos de personas. Este, en cambio, parecía ser el catálogo más visto de vídeos de explotación de menores elaborado jamás, recomendado a un público de millones de personas.

Según los estudios, la mayor parte de las personas que ven imágenes sexualizadas de niños no pasan de ahí.²⁰ Pero, en algunos consumidores, con cada vídeo aumenta la probabilidad de que se acerquen directamente a algún niño, un primer paso hacia la seducción de una víctima para una potencial agresión física. A las organizaciones para la protección de los menores les preocupaba que los vídeos pudiesen facilitar eso —o incluso que incitasen a hacerlo—, pues ponían a los niños una y otra vez ante los ojos de usuarios interesados, disipaban poco a poco sus recelos e identificaban los nombres de los familiares o las cuentas de los niños en las redes sociales.

Existía, temían los psicólogos, otro riesgo. YouTube había cultivado un enorme público de usuarios que nunca habían buscado ese contenido, sino que se habían visto arrastrados hacia él por las recomendaciones de la plataforma. Esa no era otra madriguera más. La trayectoria parecía imitar, paso a paso, un proceso que los psicólogos habían observado en reiteradas ocasiones en investigaciones sobre cómo se desarrolla la atracción por la pornografía infantil.

Para algunas personas, los impulsos pedófilos se forman en los primeros años de vida y permanecen más o menos innatos. Pero Kathryn Seigfried-Spellar y Marcus Rogers, psicólogos de la Universidad Purdue, han descubierto que los consumidores de pornografía infantil a menudo no nacieron con ese interés, sino que lo desarrollaron a lo largo de la vida.²¹ Las personas que experimentan este proceso empiezan con pornografía adulta y luego van consumiendo material cada vez más extremo siguiendo una compulsión parecida a las adicciones. A lo largo del proceso buscan contenido sexual cada vez más pervertido, es decir, pornografía un poco más tabú que la que habían visto antes. «A medida que se desensibilizan al ver esas imágenes, en caso de que lleguen a esos extremos —dijo Rogers—, van a buscar cosas que sean aún más estimulantes, aún más excitantes, aún más sexualizadas.»

Sus compulsiones las determinaba el contenido que iban

encontrando; era una suerte de entrenamiento. No era en absoluto inevitable que esas personas siguieran su ansia más allá de cualquier límite moral. Ni tampoco que esa compulsión —si en efecto la seguían— los llevase a acercarse a niños. Pero en YouTube, la segunda página web más popular del mundo, el sistema parecía haber identificado personas con esa ansia, las había conducido por una senda que las satisfacía justo al ritmo adecuado para que siguieran avanzando y luego les señalaba en una dirección muy concreta. «Esto es algo que los lleva por ese camino», dijo Rogers sobre la secuencia de vídeos de YouTube. Calificó el motor de recomendaciones como una potencial «droga de iniciación a la pornografía infantil más dura».

Investigar esta cuestión, subrayó Rogers, es un gran reto. Pocos consumidores de pornografía infantil están dispuestos a hablar de sus intereses. Algunos mienten sobre sus motivos para evitar el estigma de la pedofilia. Sería imposible hacer estudios controlados por razones éticas. Por consiguiente, los estudios como el suyo tienden a ser relativamente limitados y a menudo se centran en personas que han sido descubiertas.

Aun así, todos los expertos que Amanda y yo consultamos nos dijeron que había pruebas que hacían pensar que las conclusiones de Seigfried-Spellar y Rogers, así como sus temores acerca de su aplicación a YouTube, tenían mérito. Stephen Blumenthal, psicólogo clínico del Servicio Nacional de Salud británico que trata a personas con intereses sexuales pervertidos y que había publicado investigaciones académicas sobre sus trabajos, dijo que había visto pacientes que desarrollaban deseos pedófilos siguiendo una progresión similar. «Tiene una fuerza increíble, y las personas se sienten atraídas por ello —dijo—. Tal vez nunca habrían recorrido esa senda de no ser por las posibilidades que abrió internet.» Si bien él no tenía constancia de casos que involucrasen en concreto a YouTube, dijo, le preocupaban los paralelismos. Así como el alcance y la eficacia sin precedentes de los métodos de YouTube.

YouTube respondió con escepticismo cuando le presentamos lo que habíamos descubierto. Una portavoz calificó las preocupaciones de los psicólogos de «cuestionables». La posición de YouTube, dijo, era que la ciencia aún no había alcanzado un consenso sobre si los deseos pedófilos podían ser generados o potenciados por factores ambientales, como habían hallado Seigfried-Spellar, Rogers y

Blumenthal. La portavoz citó a una experta llamada Ethel Quayle, cuyas investigaciones —agregó— cuestionaban la existencia de cualquier efecto «puerta de entrada».

No obstante, a pesar de que YouTube embarrase su responsabilidad hablando de esa incertidumbre científica, cuando Amanda se puso en contacto con Quayle, profesora de la Universidad de Edimburgo, la experta le dijo que de hecho sus investigaciones confirmaban la teoría del efecto puerta de entrada. A los expertos, dijo, durante mucho tiempo les había preocupado que los algoritmos de aprendizaje automático creasen tales trayectorias. «El algoritmo no tiene ningún tipo de brújula moral», afirmó. En un estudio reciente sobre criminales de pornografía infantil que había realizado, aseguró, «la mayoría contaron que habían seguido enlaces, al principio de páginas pornográficas lícitas, y luego habían ido buscando material cada vez más pervertido».²²

Para las personas con una predisposición a responder a los estímulos de este tipo, dijo Quayle, las sugerencias algorítmicas como las de YouTube ofrecen «casi una ruta predeterminada de dónde van a terminar». Que la experta mencionada por YouTube creyera lo contrario de lo que YouTube había afirmado que creía me hizo dudar sobre la profundidad con que la empresa había investigado este asunto, a pesar de declarar que se tomaba en serio la seguridad de los menores. Además de eso, Quayle y otros expertos alertaban de que YouTube corría el riesgo de erosionar el tabú interno de los usuarios ante la pedofilia mostrándoles vídeos de niños junto a contenido sexual más convencional, así como enseñando el elevado número de visualizaciones de los vídeos, lo que ponía de manifiesto que tenían un gran público y, por tanto, que al parecer eran algo aceptable. Como había dicho Rogers: «Lo normalizas».

Justo después de informar a YouTube de lo que habíamos encontrado, la empresa eliminó varios de los vídeos que les habíamos enviado como ejemplos. (Otros muchos que no les mandamos, aunque formaban parte de la misma red, siguieron disponibles.) El algoritmo de la plataforma también cambió de inmediato: ya no enlazaba aquellas decenas de vídeos. Cuando preguntamos a YouTube acerca del cambio, la compañía respondió que era una coincidencia. Cuando insistí más, una portavoz me dijo que, desde un punto de vista hipotético, era posible que el cambio estuviese relacionado con

nuestras informaciones, pero que ella no podía pronunciarse ni en un sentido ni en el otro. Parecía que YouTube estaba intentando arreglar el desatino sin reconocer que hubiera que arreglar nada.

Los portavoces de la compañía nos solicitaron una gran cantidad de información: estudios, estudios alternativos, detalles sobre la metodología de los investigadores. Nos pidieron que definiésemos términos como *red* y que dijésemos a qué nos referíamos al decir que los usuarios «encontraban» vídeos. Nos preguntaron cuántas palabras tendría el artículo. Nos pedían que aclarásemos algún detalle y luego, pasados unos días, nos hacían de nuevo la misma pregunta. E insistieron en hablar con uno de nuestros editores del periódico antes de que lo publicásemos, a quien presionaron para que les desvelara informaciones sobre palabras concretas que se emplearían en el texto para describir el papel de YouTube.

Durante el tira y afloja que Amanda y yo tuvimos con YouTube, Kaiser y su equipo dieron con algo en sus análisis. Poco después de ponernos en contacto con YouTube con motivo de los vídeos de explotación de menores para notificarles que el equipo de Kaiser los había encontrado en parte rastreando la función estática del «canal relacionado» —lo mismo que había hecho en todos los experimentos anteriores a modo de comprobación metodológica de sus resultados—, la compañía eliminó esa característica del todo. Tras años en funcionamiento, había desaparecido en toda la plataforma. Cuando les preguntamos si el cambio estaba relacionado con nuestras informaciones, los representantes afirmaron desconocer que Kaiser hubiera utilizado esa función en sus investigaciones. Era una afirmación extraña, no solo porque les habíamos contado que lo estaba haciendo, sino porque YouTube había puesto objeciones a la investigación de Kaiser precisamente por este motivo.

Antes de la publicación del artículo que escribí con Amanda en el que desvelábamos lo que habíamos encontrado, la compañía me puso en contacto con Jennifer O'Connor, directora de producto de YouTube para asuntos de seguridad. O'Connor me dijo que la compañía estaba comprometida al cien por cien con la erradicación de la explotación de menores en la plataforma. Aun así, dijo, no aceptaba las investigaciones que daban a entender que el algoritmo de YouTube arrastraba a los usuarios hacia contenido más extremo, ni en este caso ni en otros. Incluso sobre una cuestión tan sensible desde una

perspectiva política y jurídica —en la que YouTube tenía unos enormes incentivos para tranquilizar al mundo de que estaba tomando todas las medidas posibles para corregirlo—, les era imposible admitir en público el rol del sistema y, por tanto, sus tendencias radicalizadoras en relación con este asunto y muchos otros.

Yo le dije que era difícil conciliar esa postura con los vídeos que habíamos visto, los millones de visualizaciones, la forma en que el sistema guiaba a los usuarios hacia ellos. Ella no aceptó ese argumento, pero insinuó que YouTube reforzaría la seguridad de todos modos. «En cuanto a los niños —dijo—, solo queremos adoptar una postura mucho más conservadora respecto a lo que recomendamos.» O'Connor no dio detalles precisos, pero sí dijo que esos elementos estaban en proceso de resolución.

Mis pensamientos volvieron a algo que había dicho Kaiser: que solo había una forma segura de impedir que eso sucediera de nuevo. En el caso de los vídeos con niños, desactivar el algoritmo. Sencillamente dejar de recomendar vídeos de niños. Eso era algo fácil dentro de las capacidades técnicas de YouTube. Tras una polémica inicial sobre los comentarios que había debajo de los vídeos de niños, la compañía había creado un sistema para identificar de forma automática esos vídeos y desactivar los comentarios.²³ Además de eso, para una compañía que factura 15.000 millones de dólares al año, ¿cuánto podía ganar con esos vídeos? Le pregunté a O'Connor si YouTube se plantearía algo así. Para mi sorpresa, dijo que la compañía «estaba avanzando en esa dirección». Prometió que pronto habría noticias.

Yo trasladé lo hablado en aquella conversación al equipo de Kaiser y también a algunas organizaciones de supervisión, que se entusiasmaron. Lo calificaron como un paso enormemente positivo, pues internet se convertiría en un lugar mucho más seguro para los niños. Pero, poco después de la publicación, YouTube «aclaró» el comentario de O'Connor. Los creadores de vídeos de YouTube dependen de las recomendaciones para impulsar el tráfico, dijeron, así que el algoritmo se mantendría activo, incluso en el caso de los vídeos de niños prepúberes.

Después de publicarse nuestro artículo, Josh Hawley, un senador republicano que solía criticar las plataformas de redes sociales, anunció un proyecto de ley como reacción a nuestros hallazgos.²⁴ Tal

como Kaiser había pedido, forzaría a YouTube y otras plataformas a desactivar las recomendaciones de vídeos de niños. Pero el proyecto de ley, de cinco páginas, que no llegó a votarse, pareció ser más bien un gesto de cara a la galería. Aun así, el senador Richard Blumenthal, un demócrata que había llevado adelante reformas serias de internet, firmó una carta con la senadora republicana Marsha Blackburn sobre este asunto que mandaron a la directora ejecutiva de YouTube.²⁵ Le hacían preguntas claras y concisas. Una en concreto se refería a si los responsables de las políticas de YouTube sobre la seguridad de los menores «participarían en las decisiones del diseño y el ciclo de vida de los productos». Ese era un problema habitual en Silicon Valley: las personas que estudian los efectos de la plataforma, al no ser ingenieros, tienen poco que decir en la fase de diseño. La carta de Blumenthal y Blackburn también repetía la pregunta principal de Kaiser: ¿por qué no se desactivan las recomendaciones de vídeos de niños?

En julio de ese año, la Comisión de Asuntos Jurídicos del Senado dedicó toda una vista (aunque hubo pocos asistentes) a esta cuestión. Su presidenta, Lindsey Graham, repitió la amenaza blandida muchas veces por Trump de eliminar las protecciones de responsabilidad de las plataformas. Blumenthal, el único demócrata que asistió, dijo que estaba «francamente decepcionado»²⁶ ante lo que consideró una no respuesta por parte de YouTube. «Esta información es indignante —dijo Hawley en la vista en referencia al artículo—. Pero creo que lo más indignante fue el hecho de que YouTube se negase a hacer algo al respecto.»

Mientras Amanda y yo estábamos elaborando el artículo, habíamos comprobado si los vídeos contenían alguna información identificativa que pudiese llevar hasta los padres. Cuando encontrábamos alguna, nos poníamos en contacto con organizaciones locales que pudieran alertar a las familias sobre lo que estaba ocurriendo y ofrecerles ayuda. Después de que una de esas organizaciones contactara con Christiane, la madre brasileña, ella se ofreció a hablar sobre su experiencia. Perpleja y enfadada, dijo que le estaba costando asimilar lo sucedido. No sabía qué contarle a su marido. Dijo que estaba confusa ante las prácticas de YouTube, incluso nos preguntó si sería posible demandar a la compañía. Y estaba preocupada porque no sabía qué hacer por su hija, cuya imagen

se había mostrado a cientos de miles de personas mediante un algoritmo que sexualizaba a los niños. ¿Cómo podía ponerla a salvo?

«Lo único que puedo hacer —decidió— es prohibirle que publique más vídeos en YouTube.»

Infodemia

NO ABRACÉIS A LOS VACUNADOS

Dos años antes de que apareciera en China un nuevo y extraño virus, un funcionario de la Organización Mundial de la Salud (OMS) llamado Andy Pattison, que es suizo, se presentó ante su jefe con un plan por si se producía precisamente un hecho como ese.¹ Las redes sociales se habían convertido en un vector de desinformación médica, le dijo Pattison a Tedros Adhanom Ghebreyesus, el director general de la OMS. Las informaciones que llegaban de trabajadores sanitarios alarmados de Brasil y otras partes habían dejado claro que las plataformas serían una gran línea de combate en cualquier emergencia de salud pública. Tenían que empezar a prepararse.

Tedros estuvo de acuerdo y creó una oficina en la sede central de la agencia en Ginebra. Como la OMS —al igual que otras agencias de las Naciones Unidas— hace buena parte de sus labores asesorando (y, si todo va bien, convenciendo) a los Gobiernos, la media docena de empleados de la oficina se centraron en cultivar las relaciones con las grandes compañías tecnológicas estadounidenses que Pattison consideraba que eran centros de poder por derecho propio. Los progresos fueron lentos. Ayudaron a Pinterest a mejorar sus resultados de búsqueda en las consultas relacionadas con las vacunas. Asesoran a Google sobre una aplicación de *fitness*. «El interés que vi fue muy mínimo. Conseguí que se reuniesen conmigo miembros de muy bajo rango de las plantillas y durante muy poco tiempo»,² ha contado Pattison. Además, según el funcionario, por lo común se tomaron pocas medidas.

Entonces, el 21 de enero de 2020, unos científicos chinos anunciaron que el virus, que había matado a cuatro personas, era contagioso entre humanos. En dos días, el número de víctimas aumentó a diecisiete. Una semana más tarde, la OMS declaró una emergencia mundial. La oficina de Pattison activó sus contactos en

Silicon Valley. «Hice ese discurso a un nivel humano: “Volved a vuestras empresas, preparaos, comenzad a crear equipos”»,³ha recordado que les dijo. Organizó llamadas entre Tedros y los jefes de Facebook, Google y otras plataformas. Mark Zuckerberg y Sheryl Sandberg sugirieron que la OMS crease páginas en WhatsApp y Facebook para publicar actualizaciones y responder preguntas de los usuarios.

El 13 de febrero, dos semanas después de que la propagación de aquel virus —que ahora se ya conocía como covid-19— fuera declarada emergencia mundial, Pattison llegó a California para participar en una reunión, celebrada en Facebook, con las principales empresas de Silicon Valley. En Estados Unidos y Europa, la vida mantenía la normalidad. Pero la percepción que Pattison tenía de la crisis ya había cambiado. Había —dijo en declaraciones a la CNBC el día de la reunión— una «infodemia».⁴Las grandes plataformas de redes sociales, entre ellas Twitter y YouTube, estaban «repletas de desinformación», aseguró.

Un mes y medio antes de que Trump trasladase esa misma afirmación desde el atril de la sala de prensa de la Casa Blanca, ya había publicaciones de Facebook que generaban cientos de miles de interacciones insistiendo en que se podía curar el virus bebiendo lejía diluida.⁵ *Influencers* de Instagram explicaban que Bill Gates había inventado el virus para justificar vacunaciones forzosas.⁶Un mensaje reenviado de WhatsApp decía que la CIA estaba haciendo acopio de provisiones.⁷Y los vídeos de YouTube que responsabilizaban de la enfermedad a las torres de telefonía 5G —dando a entender que no existía ningún virus— de repente obtuvieron millones de visualizaciones.⁸

A petición de Pattison, las compañías se comprometieron a endurecer algunas normas. YouTube eliminaría los vídeos que contraviniesen las directrices de la OMS. Facebook enviaría notificaciones a los usuarios si intentaban compartir una publicación relacionada con el covid que los moderadores hubiesen marcado como falsa. Las plataformas prometieron a Pattison y al mundo que habían aprendido la lección de los errores pasados. Esta vez harían las cosas bien.

Durante esas primeras semanas, a medida que el virus se propagaba por países enteros, con él se extendían también el miedo y

el aislamiento. Se cerraron tiendas y espacios públicos. En las grandes calles y distritos de negocios reinaba una quietud apocalíptica, desgarrada solo por el aullido de las sirenas a medida que se llenaban los hospitales; y, en algunos sitios, por la macabra presencia de camiones refrigerados delante de las saturadas morgues. Las familias se atrincheraban en casa como si se preparasen para una invasión y solo se aventuraban a salir para ir a hacer la compra, en unas tensas excursiones en las que, en nuestra ignorancia colectiva, muchos nos poníamos guantes o rociábamos con Windex productos cuyas superficies, por lo que sabíamos, podían estar fatalmente contaminadas. Las ciudades se reavivaban durante un solo minuto cada noche cuando los residentes en pisos apretujados abrían las ventanas para aplaudir con agradecimiento a los trabajadores que estaban en primera línea, aunque quizás fuera también una forma de intentar alcanzar cierto sentimiento de comunidad, para sentirse más seguros.

Durante las 23 horas y 59 minutos restantes del día, quienes sentían miedo o soledad podían acudir a esa otra ventana al mundo exterior: su ordenador. De la noche a la mañana tuvo lugar un avance de años en la adopción de las tecnologías digitales. Facebook informó de un incremento de un 70 % en el uso de la plataforma en algunos países.⁹El uso de Twitter aumentó un 23 %. Una empresa de servicios de internet calculó que el porcentaje de YouTube sobre el tráfico mundial de internet subió de un 9 a un 16 %.¹⁰El consumo de internet en conjunto aumentó un 40 %, dijo también la empresa, con lo cual en realidad el tráfico de YouTube casi se triplicó.

A pesar de los esfuerzos de Pattison por preparar a las compañías, unas cuantas etiquetas de verificación de la información no podían resolver el problema central. Las redes sociales seguían siendo una máquina diseñada para distorsionar la realidad a través de la óptica del conflicto tribal y para arrastrar a los usuarios hacia los extremos. Y la pandemia —el espectro de una amenaza invisible, omnipresente e incontrolable— activó las mismas sensaciones que alimentaban esa máquina, a una escala mayor que cualquier otro hecho desde la creación de las propias plataformas.

Era como si el mundo entero se convirtiese en una aldea afectada por el zika cuyas madres recurrieran desesperadas a los rumores virtuales, o un colectivo de hombres jóvenes solitarios que reforzasen

mutuamente su desilusión y su anomia para crear una lucha compartida contra algún enemigo inventado. Las conspiraciones sobre el coronavirus, al prometer acceso a verdades prohibidas que otros no tenían, permitían a los creyentes sentir certeza y autonomía en medio de una crisis que les había arrebatado ambas cosas. Atribuyéndolo todo a algún ser malvado o a un complot, daban a una tragedia sin sentido cierto grado de significado, aunque fuera algo misterioso. Y ofrecían a los usuarios una forma de emprender medidas, primero compartiendo su conocimiento secreto con los demás y luego contándose unos a otros que se unirían contra el culpable al que la conspiración señalase.

El relato global —que el coronavirus es un complot urdido por Ellos para controlarnos a Nosotros— estaba por todas partes ya en abril.¹¹A menudo, las conspiraciones se originaban a partir de unos usuarios diarios con muy pocos seguidores. La publicación en Facebook de un misionero de una pequeña ciudad que acusaba a Bill Gates y China de propagar el coronavirus para perjudicar a Trump.¹²El tuit de una esteticista de Houston que listaba epidemias pasadas junto con los correspondientes años electorales (las fechas estaban equivocadas) con el mensaje «El coronavirus es una enfermedad creada por los Gobiernos».¹³Un vídeo, en YouTube, de dos médicos que presentaban afirmaciones falsas de que el covid era prácticamente inocuo y que se podía prescindir de las mascarillas.¹⁴Cada publicación llegaba a un público de millones de personas, todo ello gracias a los sistemas de promoción de las plataformas.

Y este tipo de publicaciones no eran casos aislados. Los vídeos de desinformación sobre las vacunas proliferaban en YouTube diciendo a decenas de millones de espectadores que no creyeran a la «mafia médica» que pretendía incrustar microchips a sus hijos.¹⁵Facebook también experimentó un «crecimiento explosivo de las visualizaciones del contenido antivacunas»,¹⁶según un estudio de la revista *Nature*, pues el sistema de recomendaciones de la plataforma parecía redirigir un número enorme de usuarios de páginas de salud convencionales a grupos antivacunas.

A lo largo de 2020, en las redes sociales crecieron tres fuerzas de forma paralela, de las cuales las conspiraciones sobre el coronavirus no eran sino la primera. Las otras dos se demostrarían igual de

fatídicas: se trataba de cepas de extremismo virtual en aumento desde hacía mucho tiempo, con nombres que a muchos estadounidenses les habrían parecido ridículos a principios de año pero que al terminarlo se volverían aterradores; y, por otro lado, más en general entre los estadounidenses, una desinformación y una indignación ultrapartidista exageradas hasta el punto de que una rebelión armada fuera no solo algo aceptable, sino una acción, para muchos, necesaria. Estas tres fuerzas se fundamentaban en causas que existían más allá de las redes sociales: la pandemia, la reacción negativa de los blancos ante una ola de protestas por la justicia racial durante ese verano y, en especial, el presidente Trump. Pero las redes sociales impulsaron y moldearon esas causas hasta que, el 6 de enero de 2021, convergieron en un acto de violencia masiva, organizado en internet, que cambiaría la trayectoria de la democracia estadounidense, quizás para siempre.

Durante toda la primavera de ese año, mientras se propagaban las mentiras y los rumores sobre el covid, los gigantes de las redes sociales insistían en que estaban tomando todas las medidas posibles. Pero hay documentos internos que hacen pensar que en abril los ejecutivos de Facebook ya se habían dado cuenta de que sus algoritmos estaban fomentado desinformación peligrosa, que podrían haber contenido el problema de forma drástica apretando un botón y que se negaron a hacerlo por miedo a que se redujese el tráfico.¹⁷ Los investigadores de la compañía habían descubierto que las «comparticiones en serie» —las publicaciones compartidas de forma reiterada de un usuario a otro— tenían mayor probabilidad de ser desinformación. El algoritmo, al ver esas publicaciones como un buen pienso viral, potenciaba su alcance de forma artificial. Sencillamente desactivando esa ayudita, descubrieron los investigadores de Facebook, reducirían la desinformación vinculada con el covid en hasta un 38 %. Pero Zuckerberg se opuso.¹⁸ «Mark no cree que podamos generalizarlo —escribió en un memorando una empleada que había informado a Zuckerberg—. No lo introduciríamos si se produjera una compensación material en forma de MSI», añadió utilizando el acrónimo que se usa en Facebook para referirse a la participación, del inglés «meaningful social interactions» ('interacciones sociales significativas').

Ese mismo mes, los investigadores de Facebook analizaron páginas de «viralidad fabricada»,¹⁹ que republican contenido que ya es

viral para atraer seguidores, una de las herramientas favoritas de estafadores, pescadores con ciberanzuelo y traficantes de influencias rusos. Descubrieron que los charlatanes virales acaparaban el 64 % de la desinformación y el 19 % del tráfico de *todas las páginas de Facebook*: unas cifras desconcertantes, pero algo aparentemente fácil de eliminar. Sin embargo, cuando le presentaron a Zuckerberg lo que habían averiguado, él «despriorizó» su trabajo. Facebook mantuvo en secreto ambos descubrimientos.

Con las plataformas funcionando con normalidad y 6.000 personas muriendo cada día de covid —un tercio de ellas en Estados Unidos—, el mundo, ya en mayo, estaba preparado para *Plandemic* [La plandemia]. El vídeo, un falso documental de veintiséis minutos, fue subido el 4 de mayo en Facebook y YouTube. Presentaba a un científico pionero del VIH (en realidad, un exinvestigador desacreditado) que desvelaba, según el narrador, «la plaga de corrupción que pone toda vida humana en peligro». El vídeo afirmaba que el virus había sido creado para justificar la administración de vacunas peligrosas con ánimo de lucro, que las mascarillas provocaban enfermedades, que la hidroxiclороquina —un medicamento contra la malaria— servía para combatir el covid y muchas otras cosas.

La ruta del vídeo hacia la viralidad reveló —y quizás diseñó— las trayectorias de las redes sociales a través de las cuales discurriría la mayor parte del caos de 2020.²⁰ Empezando por grupos para antivacunas, conspiracionistas generales y QAnon, el relato del vídeo afirmaba todas las visiones del mundo que tenían esos grupos. Hacía aumentar la sensación que tenían esas personas de estar luchando en un gran conflicto. Y los activaba alrededor de una causa: la oposición a las sombrías fuerzas que había detrás del covid. En una semana, *Plandemic* se extendió a comunidades de medicina alternativa, luego a *influencers* de bienestar, también a páginas genéricas de yoga y estilo de vida.²¹ En paralelo, circuló por páginas y grupos que se oponían a los confinamientos, luego a páginas pro-Trump, luego a páginas para cualquier causa cultural o social que tuviese siquiera de forma vaga una tendencia conservadora. Cada grupo lo tomó como un llamamiento de la comunidad a las armas y lo entremezcló con su identidad colectiva. Muchos hicieron llegar el vídeo a Twitter e Instagram, donde volvía a empezar el proceso. En un guion que ahora

ya se estaba volviendo monótono, las compañías no reaccionaron hasta que las agencias de noticias empezaron a llamarlas para que hicieran declaraciones.²² Para entonces ya era demasiado tarde. Aunque retiraron el vídeo, las afirmaciones y llamamientos a actuar que en él se hacían habían entrado en el torrente sanguíneo digital y desde entonces han ido reapareciendo en conspiraciones de las redes sociales.

En agosto, Andy Pattison, el funcionario de la OMS, llegó a la conclusión de que tendría que cambiar de estrategia por completo con Silicon Valley. «El desafío que tengo con esas empresas es poner la sostenibilidad y la madurez por delante del beneficio absoluto —dijo a una publicación sobre desarrollo hablando con mayor claridad de lo que lo habría hecho con un medio de comunicación—. Y es duro tener una conversación así, porque todas tienen balances de cuentas que cuadrar.»²³

Durante la pandemia se reunió con regularidad con intermediarios corporativos. Pero las pruebas de que ocasionaban perjuicios en el mundo real no hicieron más que incrementarse. Según se descubrió en un estudio, aumentó la probabilidad de que los estadounidenses que usaban Facebook, Twitter o YouTube creyeran que la vitamina C podía ser un tratamiento eficaz contra el covid o que el Gobierno había creado el virus.²⁴ Millones de personas se negaban a ponerse la mascarilla o a mantener la distancia social y, meses después, se negarían a vacunarse. Los médicos decían que cada vez más pacientes rechazaban tratamientos que les salvarían la vida basándose en algo que habían visto por internet,²⁵ del mismo modo en que las familias brasileñas atormentadas por el zika lo habían hecho tan solo un año antes, y a menudo citaban las mismas conspiraciones. Trump, impulsando la indignación virtual que impregnaba su base electoral y a la vez impulsado por ella, animaba cada acción y además difundía curas falsas contra el covid y se comprometía a «liberar» los estados con medidas de confinamiento.

«Más pruebas de una plandemia»,²⁶ escribió un hombre de California en un mensaje de texto que envió a su primo en octubre, en el que le adjuntaba un enlace a un vídeo de TikTok, un tipo de intercambio que se había vuelto rutinario en Estados Unidos. «No abrazes a nadie que esté vacunado, los síntomas están relacionados directamente con la eliminación de la vacuna», respondió el primo en

referencia a una conspiración propagada por Facebook.²⁷ Los dos hombres tenían covid. Durante días fueron mandándose mensajes en los que atribuían sus síntomas a confabulaciones de los médicos o a las vacunas basándose en rumores que habían visto por internet, como se sabe por mensajes que más tarde recuperó la escritora Rachel McKibbens, hermana de uno de los hombres.²⁸ «El maldito hospital me ha empeorado las cosas. Mis pulmones no estaban en absoluto tan mal cuando llegué aquí», escribió el hermano de McKibbens. Estaba convencido de que el personal hospitalario estaba «ahí por el dinero».

A medida que su salud se deterioraba, los dos hombres buscaban pruebas de que los rumores de las redes sociales eran ciertos: uno atribuyó las fuertes menstruaciones de su esposa a las vacunas basándose en una información falsa que circulaba por Instagram. Se recomendaban el uno al otro tratamientos falsos que se habían difundido por YouTube y Facebook. Pero, por encima de todo, se instaban mutuamente a no seguir los consejos de médicos, autoridades sanitarias o incluso de McKibbens. Al empeorar los síntomas, su hermano fue hospitalizado. Pero, convencido de que los médicos lo estaban envenenando como parte de los complots que había oído una y otra vez en internet, se negó a recibir ningún tratamiento y se fue a casa, donde se atrincheró. Unos días más tarde, murió allí, solo.

SAQUEOS Y DISPAROS

A medida que en primavera y verano de ese año países enteros caían en la desinformación sanitaria, en las redes sociales se abrió una segunda madriguera paralela que atraería hombres jóvenes blancos desilusionados que buscaban un sentimiento de comunidad y una razón de ser. Se trataba de un grupo demográfico que, aunque era más reducido que el de los que sucumbían a las conspiraciones sobre el covid, resultaría ser casi igual de peligroso una vez que las plataformas hubieran hecho su trabajo. La derecha alternativa digital, antes centrada en poco más que el Gamergate o los memes de la rana Pepe, fue reclutada para unirse a un mundo de autoproclamadas milicias que se preparaban para el colapso de la sociedad, que se convencieron de que era inminente.

Para Steven Carrillo, un sargento de las Fuerzas Aéreas

estadounidenses de treinta y dos años, la evolución de chulear hablando sobre actos violentos en grupos de Facebook a la ejecución real de dichos actos empezó en marzo de 2020, cuando otro usuario le mandó este mensaje: «Empieza a planear esa operación [...]. Yo voy a dar luz verde a esa mierda».²⁹ Él respondió: «¡Pues suena bien, tío!».

Carrillo había llegado a esos grupos a raíz de una vida convulsa. De niño fue pobre; en algunos periodos vivió con sus padres en pueblos de California mientras que, en otros, con sus abuelos en una zona rural de México. En 2015 sufrió un traumatismo craneoencefálico en un accidente de tráfico, que apagó su alegre personalidad, y en 2018 perdió a su esposa por suicidio. Después de eso, entregó sus hijos a sus suegros y se fue a vivir a una furgoneta.

«Estaba absolutamente desconectado»,³⁰ ha dicho su hermana. Pero las Fuerzas Aéreas lo desplegaron a Oriente Medio durante buena parte de 2019. Unos amigos de la base lo introdujeron a grupos de Facebook que se autodenominaban milicias. En realidad, eran grupos de chat dedicados a fantasear con una insurrección o una guerra civil. Proporcionaban a sus miembros una forma de encontrarle sentido al mundo reformulándolo a partir del relato de los extremistas, en el que a una crisis individual se le da respuesta con una solución colectiva, del mismo modo en que los íncels de Reddit habían encontrado un sentimiento de comunidad imaginando una rebelión contra las feministas o los usuarios de 8chan elogiando actos de genocidio hasta que Brenton Tarrant llevó a cabo uno en Nueva Zelanda. En un guiño irónico similar, el grupo de Carrillo se había puesto el nombre de Boogaloo. Había empezado como un meme de 4chan. Sus usuarios abogaban por que hubiera una insurrección nacional con el objetivo de derribar al Gobierno y hacer realidad una utopía de derechas, invocando «Civil War 2: Electric Boogaloo», en alusión a la película de los años ochenta *Breakin' 2: Electric Boogaloo*.

En primavera de 2020, mientras las plataformas daban lugar a comunidades que temían los confinamientos por ser exactamente la toma del poder que se presagiaba en conspiraciones virales como las de *Plandemic* o QAnon, los algoritmos de las redes sociales identificaron páginas de milicias antes misteriosas como Boogaloo como el elemento exacto que atraería a esos usuarios aún más si cabe a mundos virtuales. Las publicaciones de Boogaloo se propagaron con tanta rapidez por Facebook, Twitter, Instagram y Reddit que un grupo

de monitorización del extremismo alertó de que se estaban «desarrollando insurgencias virales a la vista de todo el mundo».³¹ Los conspiracionistas del covid, entremezclados con causas milicianas gracias a los algoritmos de las plataformas, atraieron nuevos miembros a esas milicias y les ofrecieron una nueva causa urgente, mientras que las milicias dieron a los conspiracionistas una razón de ser: un conflicto inminente y definitivo con el Gobierno. Crisis y solución. Como había sucedido con movimientos similares que habían saltado de 4chan a Facebook, los grupos de Boogaloo y otras milicias cuyos miembros hacían publicaciones sobre el desencadenamiento de una guerra civil parecían usarse sobre todo para indicar pertenencia al endogrupo, en ellos no se expresaban unas intenciones reales. Pero algunos usuarios perdieron de vista la diferencia. Proliferaron las instrucciones para fabricar explosivos y armas de fuego caseras. Un archivo con el nombre «Tácticas Boogaloo» compartido en Facebook e Instagram daba instrucciones de que los «asesinatos» de «chupatintas» debían esperar hasta que sus «crímenes» fueran «demostrados ante el pueblo», pero que, de una forma más inmediata, «algunas personas tienen que caer».³² Los miembros de Boogaloo empezaron a aparecer en protestas anticonfinamientos, fuertemente armados, esperando a que comenzasen los tiros.

Durante los primeros meses de 2020, Carrillo —ahora de nuevo en California, donde había empezado a salir con una mujer de la zona — pasaba cada vez más tiempo en las páginas del grupo y fue adoptando Boogaloo como una identidad. Adornaba su perfil de Facebook con memes de la milicia y en marzo adquirió una pieza para un AR-15 en una página web que se anunciaba en grupos de Boogaloo en Facebook y prometía donar parte de sus beneficios a la familia de un presunto miembro de Boogaloo que había muerto ese mes en un enfrentamiento con la policía.³³ Algunas semanas después de su intercambio de mensajes sobre la planificación de una «operación», Carrillo se unió a grupos de Facebook y WhatsApp solo accesibles por invitación para miembros locales de Boogaloo, donde organizaban encuentros de «formación en armas de fuego» y debatían planes vagos de provocar la guerra civil que imaginaban matando agentes de policía.³⁴ En un documento de «incorporación» que rellenó en abril se le informaba, con el habitual lenguaje de videojuego de este tipo de grupos, de que «nuestras zonas de operaciones nos pueden llevar de

unos caminos de tierra al centro de una ciudad en un abrir y cerrar de ojos».

Boogaloo no era más que un grupo como otros muchos. Ese mismo mes, unos activistas por las armas de extrema derecha crearon páginas de Facebook para cada estado animando a los ciudadanos locales a protestar por los decretos que obligaban a permanecer en casa. Aunque antes de eso los activistas tenían poco seguimiento, los grupos atrajeron a 900.000 usuarios, una clara señal de promoción algorítmica.³⁵ Solo algunas decenas de personas asistieron a las protestas, algunas con fusiles del estilo de los M-15. Pero miles de personas permanecieron activas en las páginas,³⁶ que con el tiempo se fusionaron con las comunidades de crítica política y conspiraciones más mayoritarias que las plataformas seguían juntando para crear un todo más grande, lo que favorecía que se borrasen las fronteras entre esas identidades y causas.

La violencia organizada en Facebook se volvió más frecuente y más extrema. En Arkansas, en mayo, miembros de una milicia de Facebook irrumpieron con violencia en los jardines de la mansión del gobernador y quemaron una efigie del hombre. En Washington, los organizadores insinuaron que una manifestación terminaría con ataques contra ciudadanos locales que, según habían descubierto los organizadores a partir de registros policiales, habían denunciado negocios que infringían los decretos de cierre por el covid. Aunque aquellos ataques nunca se materializaron, asistieron 1.500 manifestantes, muchos de ellos armados.

Las milicias, los conspiracionistas del covid y las comunidades simpatizantes de Trump también habían empezado a adueñarse de la urgencia de vida o muerte que emanaba de otra causa con la que las plataformas los habían entrelazado una y otra vez: QAnon. Con la ayuda del algoritmo, la creencia en QAnon ahora impregnaba todas esas causas independientes,³⁷ de la misma forma en que Renée DiResta había descubierto grupos de Facebook en los que se entremezclaban los antivacunas con el Pizzagate (la conspiración que precedió a QAnon) ya en 2016.

Nina Jankowicz, experta en desinformación, dio con una trayectoria típica ese verano.³⁸ Buscaba en Facebook «salud alternativa», se unía a uno de los grupos que aparecían en las primeras posiciones y luego seguía la barra lateral con los «grupos

relacionados» generados por el algoritmo de Facebook, que con frecuencia llevaban a páginas sobre Q. Era un viaje breve y de una eficiencia tremenda que amenazaba con arrastrar a cualquiera que buscara remedios contra el covid en un momento en que la medicina aún no había encontrado ninguno. El patrón se repetía en todas las grandes plataformas, y convertía el miedo y la confusión de los estadounidenses primero en una creencia en conspiraciones más leve y luego en un QAnonismo en toda regla, un enorme impulso para que los usuarios pasaran más tiempo en las plataformas. Los canales de bienestar de YouTube y los *influencers* de *fitness* de Instagram pasaban de la astrología a las conspiraciones sobre el coronavirus y al QAnon. La red antivacunas más grande de Facebook se llenó de discursos en código sobre Q.³⁹ En TikTok proliferaron las conspiraciones sobre el Pizzagate.⁴⁰ Una *tiktoker* de veinte años, que había contribuido a desatar el resurgimiento del Pizzagate, dijo que lo había descubierto a través de un vídeo viral de YouTube. Cuando los productores de *Plandemic* publicaron una secuela, el vídeo se difundió sobre todo por páginas de Q.

Al empezar la pandemia, la causa de QAnon, con un misterio y un esoterismo de una densidad casi impenetrables, se había concentrado alrededor de una creencia central: el presidente Trump y algunos generales fieles estaban a punto de dar un glorioso golpe militar que derrocaría a la caterva que había orquestado el Pizzagate y que dominaba en secreto la vida en Estados Unidos. En la purga que vendría a continuación, el ejército ejecutaría a decenas de miles de demócratas traidores, financieros judíos, burócratas federales y liberales culturales en el National Mall. A los seguidores de Q, que en su mayoría se reunían en Facebook o YouTube y nunca accedían a los foros más duros donde aparecieron por primera vez las «gotas» de Q, se les dijo que desempeñaban un papel crucial, aunque su única labor fuera seguir las pistas y contribuir a divulgar la causa.

Ese verano, 97 creyentes confesos de QAnon se presentaron a las primarias al Congreso; 27 ganaron.⁴¹ Dos se presentaron como independientes. Los otros 25 eran candidatos republicanos a la Cámara de Representantes con la cuota al día. Los memes y referencias a QAnon predominaban sobre todo en las páginas de milicias, lo que hacía aumentar su sensación de que la violencia sería tan justificada como inevitable.

A finales de mayo, Ivan Hunter, el miembro de Boogaloo con quien Carrillo se había enviado mensajes sobre una «operación», fue con otros miembros a Mineápolis, donde tenían lugar fuertes protestas por la muerte a manos de la policía de George Floyd, un hombre negro desarmado. Los hombres se reunieron fuera de un precinto policial que había sido tomado por varios centenares de manifestantes. Hunter gritó: «Justicia para Floyd», levantó un fusil tipo Kaláshnikov y disparó trece veces contra el edificio.⁴² Nadie resultó herido. Probablemente tenía la esperanza —como los boogaloos escribían a menudo por internet— de desencadenar un enfrentamiento violento entre los manifestantes y la policía que terminase en una guerra.

Al cabo de unos días, el 28 de mayo, Carrillo publicó un vídeo de las protestas de Black Lives Matter en un grupo de Boogaloo en Facebook que frecuentaba. «Ahora está en nuestra costa, esto tiene que extenderse a todo el país [...], es una gran oportunidad para atacar a los chicos especializados de la sopa de letras»,⁴³ escribió, en alusión a las agencias estadounidenses conocidas por sus siglas —de ahí lo de «sopa de letras»—, como el FBI, pues los boogaloos creían que matar a uno de aquellos agentes podría desencadenar la guerra sin cuartel que deseaban. Y agregó: «Tenemos muchedumbres de gente enfurecida que podemos aprovechar a nuestro favor». Otro usuario local, Robert Justus, respondió: «Que empiece la fiesta». Esa noche, Carrillo le pidió matrimonio a su novia con un anillo de silicona de 25 dólares, que prometió sustituir con un diamante, y preparó una maleta para irse por la mañana.⁴⁴

Las páginas de Boogaloo como la suya no generaban una sensación de conflicto civil inminente por su cuenta. La absorbían de las redes sociales, donde esa sensación abundaba. Esa misma noche, la del 28 de mayo, Trump, impulsando ese sentimiento y a la vez bebiendo de él, publicó en Twitter y Facebook que, si las autoridades de Mineápolis no sofocaban las protestas del movimiento Black Lives Matter, él enviaría a la Guardia Nacional y haría las cosas bien. Y añadía: «Si surge cualquier dificultad, asumiremos el control, pero cuando empiecen los saqueos, también empezarán los disparos».⁴⁵ Su última expresión —que recordaba a la promesa tristemente célebre que hizo en 1967 un jefe de policía de Miami de que actuarían con determinación en los barrios afroamericanos— parecía alentar una violencia mortífera en un momento en que había mucha tensión, los

enfrentamientos en las calles iban en aumento y tanto la policía como las milicias de ultraderecha eran capaces de llevar a la práctica sus palabras.

Twitter añadió un mensaje de aviso a la publicación de Trump que decía que infringía las normas contra el «enaltamiento de la violencia» y limitó la difusión de la publicación. Pero Zuckerberg anunció que, aunque consideraba que la publicación era «divisiva e incendiaria»,⁴⁶ Facebook no la eliminaría. «Creemos que las personas deben saber si el Gobierno se está planteando desplegar agentes», explicó. Era una extraña justificación, pues amplificaba la incitación en aras del conocimiento público. Zuckerberg llamó a Trump personalmente para reiterarle las políticas de Facebook. Organizó una asamblea de toda la compañía para defender su decisión, que algunos empleados denunciaron a nivel interno.

Las tendencias de las plataformas estaban afectando a todo el mundo, no solo a los seguidores de la derecha pro-Trump. El mismo día de la publicación de Trump sobre «los saqueos y los disparos», un periodista de Mineápolis publicó unas fotos de las manifestaciones de Black Lives Matter en la ciudad. Usuarios de Twitter de izquierdas cuestionaron que las fotos podían poner a los manifestantes en riesgo de ser detenidos. Fuera pertinente o no la pregunta, la creciente indignación que le siguió la desbordó. Otros usuarios obtuvieron más atención al decir que las fotos no solo facilitaban la detención de los manifestantes, sino que garantizaban su asesinato. Citaban una conspiración desmontada desde hacía tiempo que afirmaba que los individuos fotografiados en manifestaciones de BLM tendían a morir poco después en circunstancias misteriosas.⁴⁷ La indignación creció como una bola de nieve, con decenas de miles de usuarios acusando al fotógrafo de poner en peligro adrede la vida de ciudadanos negros. Muchos expresaron el deseo de verle herido o muerto; otros prometían hacerlo ellos mismos. Algunos difundieron su número de teléfono y su dirección.⁴⁸ Esa noche, una personalidad de izquierdas de Twitter publicó fotos de la matrícula del coche de un periodista de la CNN diciendo que pertenecía a unos «provocadores mentirosos» que se infiltraban en el BLM en nombre de la policía.⁴⁹ La publicación obtuvo 62.000 interacciones y probablemente llegó a millones de personas. Ese fin de semana, unos manifestantes atacaron a un cámara de Fox News en la ciudad de Washington, mientras que otros entraron por la

fuerza en la sede central de la CNN en Atlanta. Durante todo ese verano hubo incidentes en que unos manifestantes atacaban a periodistas, aunque fueran menos frecuentes que los ataques a periodistas por parte de la policía.

La noche siguiente, el 29 de mayo, Carrillo y Justus se acercaron en una furgoneta hasta los alrededores de una protesta de BLM en Oakland. Después de dar algunas vueltas, Carrillo se subió a la parte trasera, abrió la puerta de la furgoneta y apuntó un fusil de gran potencia a dos agentes del Departamento de Seguridad Nacional que estaban fuera de un edificio de oficinas federal. Disparó contra los dos y mató a uno. Él y Justus abandonaron la furgoneta, que más adelante fue encontrada llena de armas y equipamiento para fabricar bombas.

Siguiendo la coherencia de los motivos de Carrillo, por lo visto pretendía que, por error, su ataque se viera como obra del movimiento Black Lives Matter y desencadenara más violencia. Gracias a las tendencias de las redes sociales que reafirman el partidismo, se hizo realidad al menos su primer deseo. Mientras la policía aún no había identificado públicamente a Carrillo, las páginas de Facebook pro-Trump obtuvieron cientos de miles de interacciones culpando de la muerte del agente de Oakland a los «disturbios» de Black Lives Matter, «a terroristas internos de izquierda» y «a otra ciudad demócrata mal gestionada». Describían el suceso como el último de una ola de violencia atizada por «demócratas nacionales, medios de comunicación corporativos y personal del equipo de Biden», una prolongación de la tiranía de los «confinamientos demócratas» y de unas minorías fuera de control.

A lo largo de los siguientes días, Carrillo se envió mensajes de WhatsApp con otros miembros del grupo, entre ellos Hunter, quien estaba escondido tras haber disparado contra el edificio precintado de Mineápolis, para planear más acciones violentas. Pero una semana después de su ataque, la policía siguió la pista de Carrillo hasta su casa en una población rural llamada Ben Lomond. Al llegar allí, él estaba en una colina cercana esperando para atacar. Disparó varias veces con su AR-15 casero y arrojó unos explosivos que había fabricado él mismo, con lo que mató a un agente e hirió a varios más. Había mandado mensajes a sus amigos pidiendo refuerzos, pero, tras meses de conversaciones prometedoras, lo ignoraron. Carrillo, herido por los disparos de los agentes, huyó en un coche robado, pero lo

atraparon poco después. La policía descubrió que había escrito «boog» con su propia sangre en el capó. El tiroteo había tenido lugar a menos de setenta y cinco kilómetros de las instalaciones de Facebook.

Carrillo, que no se arrepiente de nada, más adelante se casó, en una ceremonia en la cárcel, con la novia a quien le había pedido matrimonio la víspera del ataque. Su amigo Robert Justus, que también fue detenido, ha confesado a los fiscales que Carrillo le forzó a ir con él. En enero de 2022, la hermana de uno de los agentes asesinados demandó a Facebook alegando que la compañía había permitido de forma intencionada que sus algoritmos fomentasen y facilitasen el extremismo violento que había llevado a la muerte de su hermano.⁵⁰ Aunque la demanda tiene pocas probabilidades de prosperar, da cuenta de la sensación de que no puede hacerse caso omiso de la aparente complicidad de Silicon Valley en la desintegración del país a lo largo de 2020: un sentimiento que, pocas horas después del crimen de Carrillo, ya había empezado a extenderse entre los propios empleados de las compañías.

OBSTÁCULOS

El lunes 1 de junio, una semana después del asesinato de George Floyd y tres días después del ataque de Carrillo, varios centenares de empleados de Facebook colgaron mensajes en la puerta de su despacho anunciando que, durante una jornada, se negaban a trabajar. Algunos firmaron peticiones en las que se exigía más personal y cambios en las políticas por parte de sus superiores. En una primera petición, muchos condenaron en público a su empresa. «La inacción de Facebook a la hora de eliminar la publicación de Trump que incitaba a la violencia me hace avergonzar de trabajar aquí —tuiteó la ingeniera Lauren Tan—. Esto no está bien. El silencio es complicidad.»⁵¹

Nominalmente, la huelga era para protestar contra la negación de Facebook, unos días antes, de eliminar la publicación de Trump en la que amenazaba con que la Guardia Nacional dispararía contra quienes se manifestaban por la justicia racial. Pero incluso eso era, como decían muchos, solo la afrenta final. En un periodo de violencia en aumento y de desinformación sanitaria mortífera, Silicon Valley, después de contratar a los ingenieros más talentosos del mundo

prometiéndoles que contribuirían a salvar el mundo, parecía instigar o incluso provocar los problemas sociales que amenazaban con desgarrar Estados Unidos, como había hecho ya en Birmania y Sri Lanka. Como si se quisiera subrayar el argumento de los empleados, el día del paro la noticia que predominó en Facebook fue la desinformación que responsabilizaba al movimiento Black Lives Matter del doble asesinato de Carrillo.

La huelga marcó el inicio de un enfrentamiento público y de alta tensión entre los directivos de Silicon Valley y una alianza en buena medida nueva que se alzaba contra ellos: sus empleados, sus anunciantes, destacados activistas por los derechos civiles y, más adelante, dirigentes del Partido Demócrata. Eso era categóricamente distinto de hacer frente a los analistas externos o los programadores disidentes que hasta entonces habían encabezado las críticas a Silicon Valley. Era un desafío significativo para los intereses empresariales centrales del valle; un reto planteado, según sus participantes, en nombre de un mundo que cada vez más veía a esas compañías como una amenaza peligrosa.

Las preocupaciones por la desinformación electoral llevaban meses en aumento. A finales de mayo, el día posterior al asesinato de Floyd, Trump había publicado una retahíla de falsedades oponiéndose a las políticas de voto anticipado de California; había escrito: «Estas elecciones van a estar amañadas». Tras años permitiendo los exabruptos de Trump, al final Twitter había actuado; más o menos. Adjuntó un pequeño recuadro con el título «Get the facts» (Aquí encontraréis los hechos), que enlazaba a otra página que verificaba meticulosamente las afirmaciones de Trump. Se trataba, más que nada, de algo simbólico.

Tras la huelga en Facebook, para intentar apaciguar a los empleados, Zuckerberg anunció que apoyaba al movimiento BLM. No obstante, ese mismo día, la publicación más popular en su plataforma fue un vídeo de un personaje de derecha, Candace Owens, que afirmaba que «la brutalidad policial con motivaciones raciales es un mito» y que George Floyd era «un delincuente» y «un ser humano despreciable».⁵² El vídeo obtuvo 94 millones de visualizaciones, una audiencia parecida a la de la Superbowl.

A lo largo del mes de junio salieron a la luz informaciones de que Joel Kaplan, el lobista conservador en nómina de Facebook,

descafeinaba las políticas de la compañía de formas que protegían a Trump ante las normativas contra la desinformación.⁵³ Así pues, la compañía era cómplice de los esfuerzos del presidente por utilizar la plataforma como medio de presión contra los funcionarios de los estados para que estos ocultasen votos y alterasen el resultado electoral. Para muchos, eso era una prueba de que, incluso en un momento en que todo estaba en juego, no se podía confiar en que Silicon Valley hiciera lo correcto. Algunas organizaciones pro derechos civiles dirigieron una campaña («Stop Hate for Profit», ‘Basta ya de odio con ánimo de lucro’) en la que presionaron a los anunciantes para que hicieran boicot a Facebook.⁵⁴ Varias empresas y agencias de publicidad se sumaron. Algunas retiraron presupuestos publicitarios por valor de 100.000 dólares diarios.

Es improbable que eso hiciera mella en los 80.000 millones de dólares que ganaba Facebook en ingresos publicitarios anuales. Pero la amenaza para su balance de cuentas sí inspiró algo que podríamos llamar medidas, y no solo en Facebook. A finales de junio, Facebook e Instagram eliminaron Boogaloo de sus plataformas.⁵⁵ YouTube eliminó a varios destacados extremistas blancos, entre ellos el organizador de «Unir a la derecha» Richard Spencer, casi tres años después de los sucesos de Charlottesville.⁵⁶ Quizás lo más significativo fue que Reddit cerró dos mil comunidades que, según la plataforma, habían difundido discursos de odio, entre ellas The_Donald, el punto de encuentro oficioso de la extrema derecha digital.⁵⁷ Estas medidas supusieron un punto de inflexión, un reconocimiento implícito de que acoger discursos de odio era en efecto una política que favorecía su propagación. Que los discursos positivos, en esas plataformas, no vencerían a los discursos negativos de forma natural. Aun así, las eliminaciones llegaron, como tantas otras cosas, demasiado tarde. La identidad y la mentalidad sediciosa nacionalista blanca, junto con decenas de conspiraciones y falsedades que la justificaban, ya eran endémicas, estaban arraigadas en las comunidades del mundo real — más extensas— a las que las plataformas las habían inculcado con tanta eficacia.

Los críticos con Silicon Valley siguieron presionando para que se hicieran cambios de mayor calado. Un programador de Facebook, cuyo último día en la empresa coincidió que fue el 1 de julio, publicó un vídeo de despedida de veinticuatro minutos en el que, flanqueado

por unos muñecos de peluche y vestido con una serie de camisetas de Facebook de varios colores, alertaba de que «Facebook estaba perjudicando a la gente a gran escala. Estamos fracasando. Y lo que es peor: hemos incorporado este fracaso a nuestras políticas».⁵⁸

La amenaza creciente de ensuciar la reputación de la compañía a los ojos de los anunciantes fue suficiente para que Zuckerberg y Sandberg se reuniesen, a principios de julio, con los organizadores del boicot. Los dirigentes de las organizaciones civiles salieron furiosos. Declararon ante la prensa que los habían tratado con condescendencia presentándoles unas promesas vacías y habían querido manipularlos con unas afirmaciones de progreso que no parecían entender la tecnología de la compañía.⁵⁹ Sus declaraciones le sonaban familiares a cualquier periodista que hubiese entrevistado a responsables de las políticas de Facebook.

Casualmente, esa semana se publicó una auditoría independiente de las políticas y prácticas de Facebook que llevaba dos años en marcha.⁶⁰ Bajo presión pública por su rol en las elecciones de 2016 y los posteriores escándalos de privacidad, Facebook la había encargado en 2018 a un despacho de abogados especializado en derechos civiles, y desde entonces la ha blandido como una prueba de su compromiso con las mejoras. Pero los auditores, que tuvieron un acceso que Facebook había insinuado que demostraría que los críticos estaban equivocados, concluyeron que sus algoritmos fomentaban la polarización y el extremismo, sus políticas permitían que la desinformación electoral campara a sus anchas y sus prácticas internas tenían poca sensibilidad por los perjuicios en el mundo real.⁶¹ Esas acusaciones no eran nuevas ni mucho menos. Pero ahora, presentadas por los auditores que había elegido la propia empresa, fundamentadas en parte en información interna y formuladas no por unos tecnólogos disidentes o unos misteriosos académicos sino por la voz respetada de unos defensores de los derechos civiles, tenían un peso especial entre la sociedad, pues eran una confirmación autorizada de mayor relevancia que lo que se había publicado hasta entonces. El informe puso a los dirigentes de Silicon Valley a la defensiva: podían desautorizar a Guillaume Chaslot diciendo que era un ofendido o a Jonas Kaiser alegando que estaba mal informado, pero no podían hacer caso omiso de unos respetados abogados expertos en derechos civiles sin correr el riesgo de que hubiera más rebeliones de

empleados o anunciantes.

Finalmente, en agosto, Facebook y Twitter hicieron algo impensable hasta ese momento: eliminaron una publicación de Trump. El presidente había publicado un vídeo en el que se decía que los niños eran «casi inmunes» al covid. Poner en peligro la salud infantil fue, al fin, la gota que colmó el vaso. YouTube siguió sus pasos. (Las compañías tal vez tenían la esperanza de que aquello sirviera de aviso, pero no alteró el comportamiento de Trump. Dos meses y muchas mentiras después, eliminaron otra de sus publicaciones más o menos por los mismos motivos.) Mientras tanto, Trump siguió beneficiándose de la promoción algorítmica, cuyo valor excedía de largo la carga de cualquier pequeño toque de atención: sus cifras de participación en Facebook superaron las de Joe Biden a razón de cuarenta a uno ese último mes de verano, aunque Trump obtuviera peores resultados en las encuestas.⁶² una prueba más de que las plataformas no reflejaban la realidad, sino que creaban una realidad propia.

En septiembre, Estados Unidos estaba precipitándose hacia unas elecciones en las que parecía estar en juego la democracia misma. Trump y algunos de sus aliados insinuaban que intervendrían contra la votación por correo, que al parecer favorecería a los demócratas. También daban a entender que esperaban que la mayoría conservadora del Tribunal Supremo anularía las elecciones si ellos perdían. Y que se negarían a entregar el poder. Los temores a que se cometiese fraude electoral durante las elecciones y a que hubiera violencia de patrullas ciudadanas después, eran generalizados. La presidenta de la Cámara de Representantes, Nancy Pelosi, lamentando la omnipresencia de la desinformación de QAnon en Facebook, dijo: «No sé cómo pueden mirarse en el espejo el consejo de administración de Facebook o sus altos directivos. Su decisión es clara. Su plan de negocio es hacer dinero gracias a la intoxicación, y esta es la senda que han decidido recorrer».⁶³

Bajo presión, ese mes Facebook anunció que impediría a los candidatos declarar falsamente que habían ganado los comicios y eliminaría todas las publicaciones que citasen el covid para disuadir a los ciudadanos de acudir a las urnas en persona. Impuso una penalización mucho más leve a una de las tácticas favoritas de Trump.⁶⁴ «El contenido que pretenda deslegitimar el resultado de los comicios» o «la legitimidad de los métodos de votación» se marcaría

tan solo con una «etiqueta informativa». Asimismo, Facebook e Instagram no aceptarían nuevos anuncios políticos la semana antes de las elecciones, con lo cual se impediría cualquier intento de introducir contenido con el fin de manipular a los votantes en las plataformas en el último minuto. Instagram dio un paso más. Hasta que no se hubiesen celebrado las elecciones, los usuarios situados en Estados Unidos que siguiesen un *hashtag*, con independencia del tema que tratase, ya no podrían clasificar las publicaciones para ver cuáles eran más recientes. «Hacemos esto para reducir la propagación en tiempo real del contenido potencialmente perjudicial que podría aparecer en los días cercanos a las elecciones»,⁶⁵ dijeron en un comunicado de prensa. Esos dos cambios, aunque eran pequeños, cruzaban un umbral importante: limitaban las funciones básicas de los productos por el bien de la sociedad.

Un mes después de las elecciones, Twitter anunció los cambios más sustanciales que haya hecho cualquier plataforma.⁶⁶ Las cuentas con muchos seguidores —también las de políticos— estarían sujetas a unas normas más estrictas que las demás: lo contrario a las dispensaciones especiales de Facebook. Las publicaciones que infringiesen las normas se eliminarían o aparecerían detrás de un recuadro de aviso. Trump ya tenía catorce recuadros de esos, que funcionaban al mismo tiempo como elementos de verificación de la verdad y como reductores de la velocidad, pues ralentizaban la facilidad con que los usuarios podían leerlos o compartirlos. Más adelante, Twitter prohibió a los usuarios compartir o poner «me gusta» a las publicaciones de Trump que no se ajustaban a las normas. Apartados esos elementos sociales, los efectos de sus tuits parecieron disminuir en un grado considerable.

Twitter también añadió algo que habían pedido durante mucho tiempo los expertos externos: obstáculos. Por regla general, los usuarios podían compartir una publicación clicando «retuit», lo que al instante colocaba el mensaje en sus propios canales. Ahora, al darle a «retuit» se abriría un cuadro de diálogo que animaría al usuario a añadir algún mensaje de creación propia. Eso forzaba a hacer una pausa, lo que redujo la facilidad a la hora de compartir tuits. Se trataba de una intervención leve, pero sus efectos fueron significativos: los retuits descendieron un 20 % en total, dijo la compañía, y por tanto la difusión de desinformación. Twitter había

ralentizado a propósito la participación, con lo cual perjudicaba su propio interés económico, además de contravenir el dogma imperante durante décadas en Silicon Valley que insistía en que el incremento de la actividad virtual solo podía ser beneficioso. Al parecer, el resultado fue que el mundo se convirtió en un lugar menos desinformado y, por tanto, mejor.

Sin embargo, lo más sorprendente fue que Twitter desactivó de forma temporal el algoritmo que colocaba los tuits de gran viralidad en los canales de noticias de los usuarios, aunque no siguieran al autor del tuit. La compañía calificó la iniciativa para «ralentizar» la viralidad como un «sacrificio que merece la pena con vistas a fomentar una amplificación más reflexiva y explícita».⁶⁷ Por lo que yo sé, esa fue la primera vez que una gran plataforma desactivó por voluntad propia su algoritmo. Era una forma implícita de admitir justo aquello que las compañías habían evitado durante tanto tiempo: que sus productos podían ser peligrosos, que las sociedades serían más seguras si se desactivaban algunos aspectos de esos productos y que estaba a su alcance hacerlo.

Luego llegaron las medidas contra QAnon. Los bloqueos parciales que se habían impuesto ese año, con la eliminación de ciertas cuentas y grupos, no habían sido eficaces. Al fin, Facebook e Instagram impusieron bloqueos completos al movimiento en octubre, y Twitter fue suprimiendo de forma paulatina cuentas vinculadas con Q.⁶⁸ La directora ejecutiva de YouTube, Susan Wojcicki, dijo solo que YouTube eliminaría los vídeos que acusasen a una persona de participar en conspiraciones relacionadas con Q para acosarla o amenazarla.⁶⁹ Esta pequeña modificación de las normas fue el único cambio de las políticas de YouTube antes de las elecciones. Pero, como en el caso de Boogaloo y muchas otras corrientes oscuras, era demasiado tarde. Tras años en que esas comunidades extremistas habían estado madurando en las plataformas hasta llegar a ser movimientos masivos, la imposición de bloqueos solo los relegaba a plataformas más privadas, donde tenían libertad para descender aún más en el extremismo. Incluso en las plataformas mayoritarias, muchas siguieron al acecho de alguna forma. También en octubre, agentes del FBI detuvieron a varios miembros de Boogaloo que estaban acumulando armas y explosivos para el secuestro y potencial asesinato del gobernador de Míchigan. Se habían organizado en parte

en un grupo privado de Facebook.

Las características más generales de las plataformas no se modificaron. Durante las semanas previas a las elecciones, Facebook se llenó de llamamientos a la violencia contra los enemigos de Trump. Los investigadores digitales identificaron al menos 60.000 publicaciones que invocaban actos de violencia política:⁷⁰ «Los demócratas terroristas son el enemigo y deben ser asesinados todos». «La próxima vez que veamos a Schiff, debería estar colgando de la horca #muerteparaschiff.» (Adam Schiff es un congresista demócrata que encabezó el primer intento de *impeachment* contra Trump.) Otros diez mil incitaban a iniciar una insurrección armada en caso de que Biden ganase. Y hasta 2,7 millones de publicaciones hechas en grupos políticos contenían comentarios de incitación a la violencia en términos más generales como «matadlos» o «disparadles».

Era justo lo que había visto en Sri Lanka y Birmania —cada vez más consignas pidiendo sangre, explícitas y al unísono— antes de que esas sociedades se precipitaran en los propios actos de violencia que tantas veces habían amenazado con llevar a cabo. Pensé en Sudarshana Gunawardana, el ministro del Gobierno de Sri Lanka que, al ver cómo la incitación propagada en Facebook arrasaba su país, había mandado desesperado avisos a los representantes de la compañía, que no prestaban atención. Cuando por fin se acabaron los disturbios, había lamentado: «Somos una sociedad, no solo un mercado». Ahora les tocaba a los estadounidenses suplicar a Silicon Valley, con poco éxito, que recordase antes de que fuera demasiado tarde que no éramos solo un mercado para que ellos obtuvieran beneficios.

El día de las elecciones, dos candidatos Q obtuvieron escaños en el Congreso: Lauren Boebert, de Colorado, y Marjorie Taylor Greene, de Georgia. Greene también había repetido las afirmaciones de Alex Jones de que las matanzas en los colegios eran simuladas y, en un recordatorio de que la violencia política era esencial para la causa, había puesto «me gusta» a una publicación de Facebook que reclamaba que Barack Obama fuese ahorcado y a otra que decía que la portavoz de la Cámara de Representantes, Nancy Pelosi, debería «recibir una bala en la cabeza».⁷¹

El ascenso de QAnon, no obstante, quedó eclipsado ante los estadounidenses por otro suceso: Trump había perdido. Dos días

después de la votación, cuando la mayoría de los medios de comunicación aún no habían declarado formalmente un vencedor, Facebook anunció que había eliminado un grupo por desinformación electoral. Llamado «Stop the Steal» (Detengamos el robo), afirmaba que la supuesta derrota de Trump era en realidad obra de unas fuerzas misteriosas. La página obtuvo 338.000 miembros en menos de un día, por lo que fue uno de los grupos con un crecimiento más rápido de la historia de Facebook.⁷² Se llenó de conspiraciones, llamamientos a la violencia y sobre todo invocaciones a QAnon.

A pesar de todo, con Trump humillado y a poco tiempo de dejar el cargo, quizás el extremismo virtual que durante mucho tiempo había alentado también se disiparía. Al fin y al cabo, Silicon Valley siempre había dicho que sus plataformas solo reflejaban opiniones y sucesos del mundo real. Tal vez lo peor ya había pasado.

LA GRAN MENTIRA

Durante todo ese otoño, Richard Barnett, un contratista de obras de sesenta años y entusiasta de las armas de un pueblo de Arkansas, había estado cayendo por la madriguera. Obsesionado con Facebook, republicaba repetidamente conspiraciones que ya eran virales sobre el covid, las vacunas y a favor de Trump que circulaban por la plataforma: un típico superpublicador de Facebook, no muy distinto de Rolf Wassermann, el artista alemán que amplificaba lo que la plataforma le ponía delante. Pero Barnett estaba absorbiendo un ecosistema de las redes sociales que era mucho más tóxico que el de Alemania. Asistió a una concentración organizada en Facebook en el capitolio de su estado en septiembre, con un AR-15, para protestar contra las restricciones por el covid. Creía, como dijo más adelante un amigo suyo, que unos poderes misteriosos tenían la intención de aprovechar la pandemia para insertar microchips en la frente de los ciudadanos, una vaga versión de las creencias de *Plandemic* y QAnon.⁷³ Organizó un grupo de apoyo a la organización benéfica Save Our Children, de la que se había apropiado QAnon por sus labores contra el tráfico de menores.⁷⁴ En una foto de un encuentro celebrado en octubre, él y una docena de hombres llevaban fusiles de estilo militar delante de una pancarta que decía «LOS PEDÓFILOS MUERTOS NO

PUEDEN REINCIDIR», una alusión encubierta a los demócratas.

Barnett era un representante típico de la horda de usuarios de las redes sociales —algunos afiliados a milicias o a grupos de QAnon; algunos, como él, solo seguían el algoritmo— que durante todo 2020 habían estado preparándose para las grandes batallas que sus conspiraciones les contaban que eran gloriosamente inminentes. Cuando varias publicaciones virales de Facebook afirmaron que la victoria de Biden era un fraude, ellos estaban preparados para creerlo; e incluso para actuar.

«Debemos poner fin a este Gobierno corrupto»,⁷⁵ publicó un usuario tres días después de las elecciones en el grupo de una milicia en Facebook cuyas decenas de miles de miembros ya se declaraban «preparados y armados». Se trasladaron a una página privada, mejor para planificar. «Si no los eliminamos ahora, nos van a eliminar ellos a nosotros», escribió uno. Y otro respondió: «Ha llegado la hora de derribarlos».

Las milicias y los grupos de Q, pese a su influencia en las plataformas, no eran más que una pequeña facción comprometidísima de usuarios pro-Trump. Pero los sistemas de las plataformas empezaron a arrastrar cantidades mayores de usuarios a gran velocidad hacia los extremistas. Y lo hicieron promoviendo contenido que difundía la misma mentira que Trump utilizaba para intentar mantener la presidencia y que había dado vida al grupo «Stop the Steal», que fue cerrado enseguida: Trump había ganado, los demócratas habían cometido un fraude electoral masivo y los patriotas tenían que anular aquellos resultados erróneos. Eso terminó por conocerse como la Gran Mentira. Es imposible saber a ciencia cierta lo lejos que habría llegado sin las redes sociales. Pero las plataformas lo promovieron a una escala que de otro modo habría sido imposible y —quizás eso sea lo más poderoso— entrenaron a los usuarios para que se la repitieran unos a otros como una verdad apremiante. La semana posterior a las elecciones, las veinte publicaciones de Facebook con más interacciones que contenían la palabra *elecciones* estaban todas escritas por Trump.⁷⁶ Las veinte llevaban un recuadro indicando que la publicación era engañosa, lo cual al parecer tuvo poco efecto. Las publicaciones de Trump representaron veintidós de las veinticinco publicaciones con mayor interacción de Facebook en Estados Unidos.

Los rumores que validaban esa mentira se hicieron virales una y

otra vez. Biden admitía que el fraude electoral era generalizado, se les decía a los usuarios. Las papeletas demócratas de Pensilvania llevaban los nombres de personas muertas. Un canal de YouTube de derechas informó que se había descubierto a trabajadores electorales de Detroit con maletas de votos fraudulentos.⁷⁷ Otro decía a sus 1,8 millones de suscriptores que los encuestadores se habían retractado de sus anuncios de que Biden había ganado. (Cuando un periodista preguntó al operador del canal por qué YouTube no había aplicado su normativa sobre desinformación electoral contra él, respondió: «YouTube se ha portado fenomenal».)⁷⁸ Una personalidad de Twitter se inventó la información de que en Míchigan se estaban «echando a la basura» papeletas de Trump.⁷⁹ Una destacada cuenta de desinformación, @Breaking911, anunció que un empleado de correos corrupto había huido a Canadá con votos robados.⁸⁰

A mediados de noviembre, unos investigadores de Facebook hicieron un descubrimiento alarmante: un 10 % de las visualizaciones de contenido político hechas desde Estados Unidos —o un 2 % del total de visualizaciones— se correspondían con publicaciones que afirmaban que las elecciones habían sido un fraude.⁸¹ En una plataforma tan enorme —que contiene miles de millones de publicaciones diarias que abarcan todas las noticias de todos los medios de comunicación de todas las ciudades, miles de millones de grupos de debate sobre cualquier tema que uno pueda imaginar, conversaciones triviales, fotos familiares, fichas de empresas, el conjunto de la experiencia humana— era desconcertante que una sola afirmación —y más aún una afirmación peligrosamente falsa— alcanzase un porcentaje tan alto del ruido habitual.

Y la situación no era mucho mejor en YouTube. Newsmax TV, un canal que promovía con gran intensidad conspiraciones de fraude electoral por parte de los demócratas, vio dispararse sus visualizaciones de tres millones en octubre a la friolera de 133 millones en noviembre. Chaslot, rastreando una vez más las recomendaciones de YouTube, descubrió que la plataforma mostraba Newsmax a los usuarios después de que hubieran visto medios de comunicación tradicionales como la BBC, canales de izquierdas o incluso *The Ellen DeGeneres Show*. Newsmax estaba de repente en el 1 % superior de los canales de YouTube más recomendados.⁸² Chaslot identificó el mismo patrón en New Tang Dynasty TV, una cloaca de

desinformación electoral pro-Trump dirigida por el movimiento religioso Falun Gong, cuya audiencia se multiplicó por diez. Los vídeos de YouTube que difundían la Gran Mentira fueron vistos 138 millones de veces la semana posterior a las elecciones.⁸³ En comparación, 7,5 millones de personas habían visto la cobertura durante la noche electoral en todas las grandes cadenas de televisión. A principios de diciembre, unos días después de que Chaslot publicara sus resultados, YouTube anunció al fin que eliminaría las falsas afirmaciones de fraude electoral, pero, extrañamente, no retiraría los vídeos falsos que ya estaban colgados, y solo castigaría la infracción de esta norma a partir del 20 de enero.

La plataforma continuó saturada de mentiras sobre las elecciones.⁸⁴ Cualquier persona que leía noticias en las redes sociales tenía motivos para llegar a la conclusión —como hizo Richard Barnett, el superusuario de Arkansas— de que había «un montón de pruebas» que demostraban que Trump había ganado los comicios.⁸⁵ Barnett amplificó la conspiración a sus amigos y activó su indignación del mismo modo en que la mentira había activado la suya; actuaba, así, como un diligente operario de la máquina de las redes sociales.

El 19 de diciembre, un mes y medio después de las elecciones, Trump tuiteó: «Gran protesta en la ciudad de Washington el 6 de enero. ¡Yo estaré, va a ser una pasada!». Ese día estaba previsto que el Congreso certificase la victoria de Biden. Trump estaba presionando a los congresistas para que desechasen votos de Biden, con lo cual anularían su victoria en lo que sería un verdadero golpe de Estado. Algunos legisladores republicanos ya daban señales de que estarían dispuestos a hacerlo, y en última instancia muchos lo hicieron. Quizás una concentración en la Casa Blanca, pensaban Trump y sus aliados, podía presionar a los que se resistían.

En internet, muchos de los seguidores de Trump se tomaron su mensaje como una validación de todo lo que llevaban meses contándose unos a otros. La caterva de demócratas abusadores de niños estaba a punto de ser desenmascarada y probablemente ejecutada. Trump llamaría al ejército, y a las milicias temerosas de Dios más les valía estar allí para apoyarle. «¡Los #patriotas deben ser tan violentos como BLM/Antifa! ¡¡Este es nuestro espíritu!!», escribió un miembro de los Proud Boys, un grupo nacionalista blanco, en Parler, un clon de Twitter que había crecido gracias a la extrema

derecha después de que bloqueasen al movimiento en las plataformas convencionales. En TheDonald, un portal de chat que imitaba la subsección de Reddit ahora eliminada, más de un 80 % de los debates sobre el acto del 6 de enero contenían llamamientos explícitos a la violencia.⁸⁶ Algunos usuarios publicaron planos del edificio del Capitolio en los que señalaban túneles y entradas.

A finales de diciembre, muchos usuarios ya iban poniéndose de acuerdo en un plan. Algunos traerían armas y explosivos a Washington. Otros instigarían a una multitud lo bastante numerosa para sobrepasar a la policía del Capitolio. Entrarían por la fuerza en el edificio y detendrían la certificación de las elecciones. Y luego, como escribió un usuario de Parler: «Vamos a destrozar el Congreso».⁸⁷ Todo un universo de grupos de Facebook, alzándose con una sola voz, promovió la concentración de Trump como la gran batalla para la que se habían estado preparando. Memes que parecían folletos, omnipresentes en la plataforma, instaban a asistir, a menudo con eslóganes de milicias que indicaban el inicio de una rebelión armada. Muchos incluían lemas de QAnon que lo calificaban como «the storm» ('la tormenta' o también 'el asalto'), la purga sangrienta que Q había presagiado. Y llevaban un *hashtag* que recordaba los planes de los foros de extrema derecha a los que enlazaban muchos de los grupos: «#OccupyCongress» ('Ocupar el Congreso').

«¡¡¡Este es NUESTRO PAÍS!!!»,⁸⁸ escribió Barnett en su página de Facebook unos días antes de Navidad, con lo que insistía a sus seguidores que fueran con él a la concentración. Ese mismo día, publicó una foto en la que aparecía con un fusil, junto con un texto que decía que había llegado a este mundo pateando y chillando, cubierto de sangre, y que estaba dispuesto a irse de la misma forma. Al cabo de unos días, el 2 de enero, llevó un cartel que decía «EUA República Bananera» en una concentración de «Stop the Steal» en su pueblo de Arkansas, donde le contó a un periodista sus ambiciones de restablecer la supuesta victoria de Trump. «Si no te gusta, manda a alguien para que venga a por mí, pero yo no me voy a rendir fácilmente.»⁸⁹

Como había sucedido antes en muchos casos —tanto con los incels como con los boogaloos—, lo que había empezado como unas provocaciones chulescas en internet para encontrar un sentimiento de comunidad en medio de la desorientación terminó convirtiéndose, en

unas plataformas que recompensaban el aumento de la tensión y creaban una falsa sensación de consenso en torno a opiniones de lo más extremas, en una sincera voluntad de pasar a la acción. «Hoy he tenido aquella conversación tan difícil con mis hijos: que podría ser que papá no volviese de Washington»,⁹⁰ escribió un usuario de TheDonald el día antes de la concentración. Su mensaje, que recibió 3.800 votos favorables, decía que tenía la intención de cumplir su juramento militar de «defender a mi país de todos los enemigos, tanto extranjeros como internos».

El foro se llenó de historias como esas. Eran mensajes de potenciales mártires, que se correspondían casi palabra por palabra con los que hacían, en vídeos y publicaciones en las redes sociales, los terroristas suicidas yihadistas el día antes de cometer un atentado. «Hoy me he despedido de mi madre. Le he dicho que he tenido una buena vida»,⁹¹ escribió otra usuaria. «Si nuestros “dirigentes” no hacen lo que toca y tenemos que asaltar el Capitolio, yo voy a hacerlo. Nos vemos allí, *pedes*»,⁹² escribía la mujer, despidiéndose con una referencia a un sobrenombre inspirado en YouTube y Reddit que había creado lazos entre la comunidad durante años.⁹³ «Será el honor de mi vida luchar a tu lado.»⁹⁴ Los usuarios publicaron decenas de respuestas de apoyo: «Esos edificios nos pertenecen [...]». «Lo que deberían asaltarse son las cámaras del Congreso si se llega a tal extremo, no la Casa Blanca [...]». «Traed madera, construid la horca fuera del Congreso, estad preparados mentalmente para sacarlos y colgarlos [...]». «Yo llevaré un arma a la vista y mis amigos también. No hay suficientes policías en Washington para detener lo que se avecina.»

Barnett llegó temprano a los terrenos de la Casa Blanca el 6 de enero. Mientras esperaba a que apareciese Trump, sacó el móvil y publicó un vídeo en Facebook. «Ya estamos llegando todos —dijo—. Preparaos para la fiesta.» Sobre el mediodía, al sur de la Casa Blanca, Trump empezó su discurso. Cuando se acercaba al final, dijo a la multitud: «Vamos a ir hasta el Capitolio», dando a entender que ese sería el siguiente paso en su objetivo, declarado desde hacía semanas, de impedir que el Congreso certificase los votos electorales, como estaba haciendo en ese momento. «Porque nunca vais a recuperar nuestro país con una actitud débil. Tenéis que demostrar fuerza, debéis ser fuertes», añadió Trump, ante lo cual la multitud rugió en aprobación. Miles de personas se dirigieron entonces al Capitolio,

aunque sin Trump, quien, a pesar de asegurar a la multitud «estaré allí con vosotros», había regresado a la Casa Blanca para seguir los acontecimientos por televisión.

El otro líder de la insurrección —a fin de cuentas, quizás su líder real— ya estaba sobre el terreno, metido en los bolsillos de todos los participantes que llevaban un móvil. El 6 de enero fue la culminación del trumpismo, sí, pero también la de un movimiento construido en las redes sociales y por ellas. Fue un acto que se había planificado, con días de antelación, sin planificadores. Había sido coordinado por miles de personas sin coordinadores. Y ahora sería ejecutado por medio de una voluntad colectiva guiada digitalmente. A medida que la gente llegaba al Capitolio, encontraban otros manifestantes que habían llegado antes que ya estaban metiéndose con los pocos policías que montaban guardia. En los terrenos de los alrededores se había construido una horca de madera en la que colgaba una soga. A dos bloques de allí, la policía descubrió una bomba casera en la sede central del Comité Nacional Republicano. Luego encontraron otra en el Comité Nacional Demócrata.

La presencia policial seguiría siendo escasa durante buena parte del día. Los agentes federales y de la ciudad de Washington, pese a que fueron alertados de la naturaleza de los debates virtuales, sencillamente no vieron la concentración trumpista como la amenaza que era. Incluso una vez estalló la violencia, su respuesta se retrasó durante horas de disfunción burocrática, según se constató en un informe del inspector general.⁹⁵ Varios altos oficiales dimitieron más adelante avergonzados.

Mientras los manifestantes se unían a los congregados que empujaban las vallas, debieron hacer un cálculo sencillo del panorama. Eran miles de personas: furiosos, gritando, muchos con cascos y equipamiento de excedentes militares. El único obstáculo que había, en algunas entradas, eran unas barricadas de metal que pesaban menos de diez kilos controladas por tres o cuatro agentes con gorras de béisbol tras las cuales se abrían los pasillos para entrar en el Capitolio. Una hora después de la exhortación de Trump, la multitud superó por la fuerza el perímetro exterior. Transcurrida otra hora, lograron entrar en el edificio. Los políticos que había dentro, ocupados con la certificación de los votos, ni se imaginaban que hubiera ningún problema hasta que la policía entró de golpe y montó barricadas en

puertas internas.

«¡Estamos dentro, estamos dentro! ¡Derrick Evans está en el Capitolio!»⁹⁶ Evans, un diputado del estado de Virginia Occidental, dijo eso gritando a su móvil, mientras retransmitía en directo por Facebook, donde llevaba días haciendo publicaciones sobre la concentración.⁹⁷ En casi cualquier foto del asalto al Capitolio, se ven asaltantes con el móvil en alto. Están tuiteando o retransmitiendo en directo por Instagram, Facebook y YouTube. Eso era, como la matanza de Christchurch un año antes o los asesinatos íncels el año anterior, una escenificación, todo hecho para las redes sociales y en las redes sociales. Hasta tal punto era un producto de la web social que muchos de sus participantes no distinguían entre las vidas que vivían en el mundo digital y la insurrección que estaban llevando a cabo en el mundo real como una extensión de las identidades creadas por esas plataformas.

Muchos de los que entraron por la fuerza al edificio llevaban camisetas y gorras de QAnon. Jake Angeli, un obseso de las redes sociales de treinta y dos años, se había pintado la cara con la bandera estadounidense, se había disfrazado con unos cuernos y un tocado de piel de animal y se autodenominaba «chamán de Q». Otras personas llevaban indumentaria de camuflaje de estilo militar con parches con los nombres de milicias de Facebook. Uno se escondía debajo de una máscara gigante de la rana Pepe. «Hemos empujado y empujado y empujado, y hemos gritado “vamos” y “al ataque”»,⁹⁸ narró una florista de treinta y seis años del oeste de Texas, Jenny Cudd, en un directo de Facebook mientras entraba por los pasillos. «Hemos subido hasta lo alto del Capitolio y había una puerta abierta y nos hemos metido dentro —siguió diciendo—. Hemos derribado la puerta del despacho de Nancy Pelosi y alguien ha robado su mazo y se ha hecho una foto en la silla dedicando una peineta a la cámara.»

El hombre que había visto Cudd era Richard Barnett, el conspiracionista de Arkansas. Y la foto que se hizo —con los pies encima de la mesa de Pelosi, con chaqueta de franela y vaqueros, con los brazos extendidos en aparente éxtasis y un móvil en la mano— se convirtió, antes siquiera de que terminara el asalto, en un símbolo de la irrealidad y la humillación del día. Barnett, sonriéndonos a todos, se convirtió en el rostro de algo grotesco de la vida estadounidense, algo cuya fuerza muchos de nosotros no habíamos entendido hasta la

fecha, cuando desfiló por los pasillos del poder y se echó a la bartola dentro del Capitolio.

Algunos habían acudido para algo más que retransmitir en directo. Un grupo de ocho hombres y una mujer, con ropa de estilo militar encima de sus camisetas de los Oath Keepers, avanzaron por entre la multitud y entraron en el Capitolio. Estaban en contacto por Zello, una aplicación de pulsar para hablar, y por Facebook, la plataforma en la que su milicia se había formado y había reclutado miembros. «Estáis ejecutando una detención civil, en calidad de ciudadanos. Detened esa asamblea, tenemos motivos plausibles de actos de traición, de fraude electoral»,⁹⁹ dijo uno de ellos. «Dentro», escribió uno en Facebook. «Todos los legisladores están abajo en los túneles, tres pisos más abajo», respondió otro. Un tercero dio una orden: «Todos los miembros están en los túneles debajo de la capital, encerradlos dentro. Encended el gas».

Otros compartían las intenciones de los Oath Keepers. Peter Stager, un hombre de cuarenta y un años de Arkansas, dijo en un vídeo de Twitter: «Todas las personas que hay ahí son traidores traicionadores» y «La muerte es el único remedio para lo que hay en este edificio».¹⁰⁰ Portó una bandera estadounidense mientras subía por las escaleras del Capitolio, encontró a un policía que la muchedumbre había tumbado y le golpeó con el asta de la bandera. Angeli, el «chamán de Q», dejó una nota para Pence en su escritorio en el Senado: «Es solo cuestión de tiempo, la justicia está llegando». Pence se había escapado de los asaltantes por pocos metros. Más tarde salieron diputados diciendo que se habían apiñado en despachos secundarios o detrás de puertas cerradas, temiendo por su vida mientras las turbas enfurecidas se paseaban por los pasillos. Dominic Pezzola, un miembro de los Proud Boys que más tarde fue acusado de entrar en el Capitolio rompiendo una ventana con un escudo de un policía, «dijo que habría matado a cualquiera con quien se hubiese cruzado, también a Nancy Pelosi»,¹⁰¹ afirmó un informador del FBI. El informador dijo que Pezzola y sus amigos planeaban ir a Washington para la toma de posesión y «matar todos los “hijos de p.” que pudieran». Antiguos amigos suyos contaron a *Vice News* que Pezzola se había vuelto hiperactivo en Facebook; uno describió su evolución diciendo que había caído por una «madriguera» de las redes sociales.¹⁰²

Durante la insurrección murieron cinco personas. Ninguna de ellas era un objetivo concreto; las mató la violencia bestial de la muchedumbre descontrolada. Brian Sicknick, un policía de cuarenta y dos años superado por la multitud y rociado dos veces con spray de pimienta, regresó al cuartel general, se desmayó y murió al día siguiente. Un médico atribuyó el infarto que lo mató a causas naturales agravadas por la violencia.

La mayoría de los otros fallecidos eran miembros de la multitud; como muchos otros, arrastrados a causas extremistas, al mismo tiempo participantes y víctimas. Kevin Greeson, un usuario de Parler de cincuenta y cinco años y seguidor de los Proud Boys de Alabama — que había publicado: «¡Cargad las armas y salid a la calle!»— murió de un ataque al corazón en medio del gentío.¹⁰³ Benjamin Philips, un hombre de cincuenta años de Pensilvania que había puesto en funcionamiento una red social alternativa llamada Trumparoo, murió de un infarto. Roseanne Boyland, una mujer de Georgia de treinta y cuatro años obsesionada con Facebook, se desmayó entre la muchedumbre en un túnel del Capitolio y murió. La familia de Boyland dijo que había descubierto la extrema derecha digital y QAnon mientras buscaba un significado a su vida tras batallar durante años con la adicción.¹⁰⁴

Y, por otra parte, estaba Ashli Babbitt, una californiana de treinta y cinco años, veterana de las Fuerzas Aéreas y propietaria de una empresa de equipamiento para piscinas que había reorientado su vida en torno a QAnon y tuiteaba más de cincuenta veces al día. Babbitt, con una bandera trumpista a modo de capa, se coló a la fuerza por una ventana rota de una puerta donde había una barricada policial y, a pocos metros de los políticos que su comunidad había insistido en que había que matar, recibió el disparo de un agente de policía del Capitolio. Murió al instante. Como muchas otras personas de la multitud, John Sullivan, un usuario muy asiduo de las redes sociales, filmó la muerte de Babbitt.¹⁰⁵ «¡Eso me conmovió, me calentó!»,¹⁰⁶ dijo Sullivan sobre la grabación. Mientras Babbitt caía desplomada entre la multitud con sangre saliéndole por la boca, Sullivan dijo asombrado: «Tío, esta mierda se va a hacer viral».

En el césped del Capitolio, Matthew Rosenberg, un periodista del *New York Times* que había acudido enseguida para informar sobre aquel caos, se topó con Barnett, que ya era conocido por la foto que se

había hecho en el despacho de Pelosi tan solo unos minutos antes. Barnett había ido deambulando hasta el césped, con la camisa ahora rasgada por alguna melé del Capitolio. Llevaba en las manos una carta del escritorio de Pelosi. Hablando sin reservas con Rosenberg, Barnett le enseñó la carta de Pelosi con orgullo y le dijo: «Le he escrito una nota cruel, he puesto los pies encima de la mesa y me he rascado las pelotas». Entonces volvió a adentrarse en la muchedumbre que ahora merodeaba de forma caótica por delante del edificio saqueado del Capitolio, bebiendo cerveza y ondeando banderas, con lo que proyectaban un aire de victoria.

Después de eso, con cientos de detenidos (entre ellos Barnett) y con Trump destituido por segunda vez, las ondas expansivas seguirían reverberando durante meses. Pero antes incluso de que terminara el asalto, las reverberaciones ya habían llegado hasta Silicon Valley. «¿Podemos ver un poco de valentía y medidas reales por parte de los directivos como respuesta a ese comportamiento? —escribió un empleado de Facebook en el tablón de mensajes interno de la compañía mientras se producían los disturbios—. Vuestro silencio es, al menos, decepcionante y, en el peor de los casos, criminal.»¹⁰⁶

En una retahíla de publicaciones internas, varios empleados mostraron su enfado por la decisión de Facebook de no eliminar una publicación de Trump, escrita durante el asalto, que alentaba de forma implícita a seguir adelante: «Estas son las cosas y los hechos que ocurren», escribió, cuando «se arrebató» una victoria «sagrada y aplastante a unos grandes patriotas». La relación de la compañía con sus cerca de 50.000 empleados ya estaba en horas bajas desde los conflictos del año anterior. Ahora, mientras Facebook permitía a Trump utilizar la plataforma para atizar aquella insurrección, se derramó la frustración. «Debemos eliminarle la cuenta ahora mismo. No es momento de medias tintas»,¹⁰⁷ escribió otro empleado. Pero, en lugar de actuar contra Trump, la compañía congeló los comentarios de los debates internos.

Al día siguiente, el sindicato de trabajadores de Alphabet, que se había formado esa misma semana, publicó un comunicado condenando la inacción de su empresa. (Alphabet es la compañía matriz de Google y YouTube.) «Las redes sociales han envalentonado a un movimiento fascista creciente en Estados Unidos y somos muy conscientes de que YouTube, un producto de Alphabet, ha

desempeñado un papel clave en esta amenaza al alza», escribieron. Los trabajadores de Alphabet, añadían, habían alertado a los ejecutivos en repetidas ocasiones del rol de YouTube en «el odio, el acoso, la discriminación y la radicalización», pero estos habían «ignorado o hecho concesiones simbólicas». Instaban a la compañía, por lo menos, a «obligar a Donald Trump a ceñirse a las normas de la plataforma».¹⁰⁸

Había, si no un cambio radical en Silicon Valley, al menos un atisbo de reconsideración. Chris Sacca, uno de los primeros inversores de Twitter, tuiteó: «Tenéis las manos manchadas de sangre, Jack y Zuck. Durante cuatro años habéis racionalizado este terrorismo. Instigar una sedición violenta no es un ejercicio de libertad de expresión. Si trabajáis en estas empresas, también depende de vosotros. Eliminad esos discursos».¹⁰⁹ Al preguntarle por qué también había regañado a los empleados, Sacca respondió: «Con toda sinceridad, las únicas personas a las que escuchan son los empleados que los rodean. En el sector tecnológico, si pierdas talento, pierdes poder».

El día posterior a los disturbios, Facebook anunció que impediría a Trump utilizar sus servicios al menos hasta la toma de posesión, dos semanas después. Al día siguiente, mientras Trump seguía tuiteando en apoyo a los asaltantes, Twitter también le desactivó el perfil. YouTube, el último gran resistente, hizo lo propio cuatro días después. La mayoría de los expertos y buena parte de la sociedad estuvo de acuerdo en que bloquear a Trump era necesario pero llegaba tarde. Aun así, había una innegable incomodidad por el hecho de que esa decisión estuviera en manos de unos cuantos ejecutivos de Silicon Valley. Y no solo porque eran unos actores empresariales que nadie había elegido, sino porque, para empezar, las decisiones de esos mismos ejecutivos habían contribuido a llevar la crisis de las redes sociales hasta ese punto. Tras años en que el sector había sido conciliador con Trump y los republicanos, el bloqueo se vio de forma generalizada como una medida adoptada por interés propio. Al fin y al cabo, se había tomado tres días después de que los demócratas asumieran el control del Senado, que también controlaban la Cámara de Representantes y la Casa Blanca.

Los demócratas, cuya indignación no dejaba de acrecentarse, vieron el bloqueo como una medida superficial que se aplicó solo

cuando la salida de Trump del poder ya era inevitable. Desde la campaña de Biden habían escrito en privado a Facebook durante las elecciones expresando primero preocupación y luego indignación por lo que consideraban una actitud de pasividad por parte de la compañía. «Como ya sabréis, nunca he sido muy fan de Facebook», había dicho Biden antes al consejo editorial del *New York Times*, dando a entender que su Administración podría revocar ciertas protecciones legales de que gozaban las plataformas de redes sociales. «Nunca he sido un gran fan de Zuckerberg. Creo que es un verdadero problema.»¹¹⁰

El día después de la investidura de Biden, dos congresistas demócratas enviaron cartas a los directores ejecutivos de Facebook, Google, YouTube y Twitter. Se trataba de Tom Malinowski, que había sido un enérgico responsable de derechos humanos en el Departamento de Estado, y Anna Eshoo, que había representado al distrito de California, donde se encuentra Silicon Valley, desde 1993. Los dirigentes demócratas habían defendido Silicon Valley durante décadas. Ahora los propios congresistas de las compañías habían coescrito una carta diciéndoles: «Tal vez no haya ninguna otra entidad que acarree mayor responsabilidad por la divulgación de peligrosas teorías conspirativas a gran escala o por avivar las reivindicaciones antigobierno que las que ustedes crearon y que dirigen hoy en día».¹¹¹

Las cartas atribuían la mayor parte de la responsabilidad de la insurrección a las compañías. «El problema fundamental —escribieron a los directores ejecutivos de Google y YouTube— es que YouTube, al igual que otras plataformas de redes sociales, clasifica, presenta y recomienda información a los usuarios con que les nutre de contenido basado en la indignación, la ansiedad y el miedo.»¹¹² Las cartas dirigidas a Facebook y Twitter eran parecidas. Todas exigían cambios de gran envergadura en las políticas y terminaban con la misma advertencia: que las compañías «iniciasen una reevaluación exhaustiva de su sistema, en el que se usaba la maximización de la participación como base de la clasificación y recomendación algorítmica». El lenguaje señalaba a propósito que los demócratas habían recogido la opinión defendida desde hacía tiempo por investigadores, sociólogos y disidentes de Silicon Valley: que los peligros de las redes sociales no se resuelven solo moderando mejor o modificando las políticas. Están arraigadas en la naturaleza fundamental de las plataformas. Y son lo

bastante graves como para amenazar la propia democracia estadounidense.

Ese enero hubo otro cambio: QAnon desapareció casi por completo. «Hicimos todo lo posible. Ahora debemos ir con la cabeza bien alta y volver a nuestras vidas lo mejor que podamos»,¹¹³ publicó Ron Watkins, el administrador de 8chan (ahora rebautizado «8kun»), que mucha gente sospecha que escribió el material de Q, la mañana de la investidura de Biden. En Telegram —una aplicación social que había ganado popularidad entre los seguidores de QAnon a medida que Twitter fue poniéndoles más obstáculos— instó a los seguidores a respetar la legitimidad de Biden. Agregó: «Ahora que comienza el mandato de la siguiente Administración, acordaos, por favor, de todos los amigos y todos los recuerdos agradables que hemos vivido juntos a lo largo de los últimos años».¹¹⁴ Watkins decía al movimiento —que se sospechaba que agrupaba a millones de personas, todas ellas preparándose para una batalla final contra los malvados responsables de todos los males de su vida— que depusieran las armas. Después de ese mensaje, las publicaciones de Q, que ya guardaba un misterioso silencio desde el 8 de diciembre, se terminaron.

La sensación de que algo se había terminado era omnipresente. Un moderador de 8kun eliminó los archivos de «QResearch» ('Investigación de Q') del portal, y escribió: «Solo estoy practicando la eutanasia a algo que he querido muchísimo».¹¹⁵ Algunos usuarios empezaron a publicar sus mensajes de despedida. Otros intentaban asimilarlo: «Moderadores por favor explicad por qué Biden aún no ha sido arrestado».¹¹⁶ Uno comparaba el hecho de ver la investidura de Biden con «ser un niño y ver el gran regalo debajo del árbol de Navidad [...] pero abrirlo y darte cuenta de que era un pedazo de carbón».¹¹⁷ Sin unas plataformas mayoritarias que acelerasen su causa o la interconectasen con la web social en general, los creyentes que quedaban tenían pocos sitios a los que aplicar sus energías, intensas unos días antes. Daban vueltas y más vueltas buscando una validación que nunca llegaba, ansiando una resolución a la crisis psicológica que les habían provocado años de radicalización.

«TODO ocurrirá en los próximos cuarenta y cinco minutos»,¹¹⁸ escribió un usuario en un foro de Q durante la jura del cargo por parte de Biden. Los demócratas que hay en el escenario —prometía el usuario— «¡serán detenidos en directo por televisión con

decenas de millones observándolo asombrados!». Sería «¡el mejor día desde el Día D» y «Estados Unidos será un país unido en la celebración!». Cuando la investidura siguió adelante con normalidad, otro usuario le preguntó si se encontraba bien. Él insistió en que la victoria aún estaba por llegar y que, con ella, habría un retorno a la vida que le habían arrebatado. «He perdido amigos y una novia durante el pasado año porque se negaron a ver la verdad, ahora por fin me estoy redimiendo —escribió—. Pronto todos volverán a pedirme disculpas, este es el día más feliz de mi vida.»¹¹⁹

EPÍLOGO

Denunciar las malas prácticas

En las semanas posteriores al asalto al Capitolio se abrió una ventana. A diferencia de los periodos de críticas de 2016 y 2018, que no habían dado frutos, ahora buena parte de la sociedad por fin entendía las consecuencias de las redes sociales. La renovación de arriba abajo del concepto de las plataformas se veía como algo más necesario que nunca y, dado que gran parte de la propia Silicon Valley estaba horrorizada, quizás era incluso posible. Los expertos y los redactores sobre tecnología estaban entusiasmados con las nuevas posibilidades. Quizás unos servicios de suscripción, en los que los usuarios pagasen una cuota mensual para conectarse, podrían acabar con la adicción del sector a los ingresos publicitarios y, por tanto, a la participación. Quizás una legislación sobre responsabilidad más restrictiva podría reorientar sus incentivos.

Pero la ventana se cerró enseguida. Los gigantes de las redes sociales estaban demasiado comprometidos con el modelo financiero e ideológico vigente para acometer una transformación tan radical. Por encima de todo, se fundamentaban en los métodos que mejor conocían: una tecnología automatizada y la moderación de contenido a gran escala. Twitter, por ejemplo, incrementó los «obstáculos» añadiendo mensajes y cuadros de diálogo («¿Quieres leer primero el artículo?») para impedir que los usuarios compartieran publicaciones de un modo compulsivo. Fue un cambio significativo, pero se quedaba lejos de las propuestas hechas por Jack Dorsey de repensar la mecánica subyacente a la plataforma, que permaneció intacta. Más adelante, en noviembre de 2021, Dorsey dimitió del cargo de director ejecutivo, y se frustraron sus ambiciosas promesas.

Algunas semanas después de la insurrección, Zuckerberg anunció que el sistema de recomendaciones de Facebook dejaría de promover grupos políticos:¹ una decisión potencialmente histórica con vistas a mitigar uno de los elementos más peligrosos de la plataforma. No obstante, en una señal de que las reformas de la compañía reflejaban

concesiones a la presión externa hechas a regañadientes y no un cambio sincero de mentalidad, era la tercera vez que Facebook anunciaba ese cambio. Las promesas anteriores, en octubre de 2020 y el 6 de enero, habían sido en su mayor parte vacías, como averiguaron investigadores independientes.² Esta última promesa solo terminó cumpliéndose por la presión de los congresistas demócratas.³

Los cambios de rumbo de Facebook habían empezado incluso antes de la insurrección. Durante las elecciones, la empresa había modificado su algoritmo para promover medios de comunicación «autorizados» antes que enlaces ultrapartidistas. En diciembre, dio marcha atrás y restableció las preferencias originales de los algoritmos.⁴ Ese mismo mes, Facebook anunció también que levantaría su prohibición, impuesta poco antes, a los anuncios políticos, pero solo en Georgia, donde dos elecciones inminentes al Senado determinarían el control de la cámara. La compañía afirmaba que hacía caso a unos «expertos» anónimos que al parecer habían calificado los anuncios políticos como algo crucial para «expresar la opinión».⁵ En realidad, las organizaciones prodemocracia se oponían en buena medida al cambio. El efecto fue inmediato. Antes de que Facebook levantara la prohibición de los anuncios políticos en Georgia, en los canales de noticias de los usuarios aparecían sobre todo artículos del *Wall Street Journal* y del *Atlanta Journal-Constitution*. Al día siguiente de que Facebook cambiara de política, los sustituyeron artículos de *Breitbart*, contenido partidista y anuncios políticos.⁶ Después del 6 de enero, los críticos presionaron a Facebook para que reintrodujera la modificación del algoritmo y la prohibición de los anuncios políticos. A fin de cuentas, al principio la compañía había presentado ambas políticas como algo esencial para proteger la democracia. ¿Cómo podía dar marcha atrás? Pero el algoritmo se mantuvo en su forma más peligrosa y, ese marzo, Facebook levantó la prohibición de los anuncios políticos en todo Estados Unidos.⁷

Tan deprisa como dieron marcha atrás en las reformas, los directivos de Silicon Valley también empezaron a cuestionar hasta qué punto era responsabilidad suya cambiar las cosas. «Creo que esos hechos se organizaron en buena medida en unas plataformas que no tienen nuestras capacidades para detener el odio, no tienen nuestros estándares y no tienen nuestra transparencia»,⁸ dijo sobre la insurrección del 6 de enero Sheryl Sandberg, número dos de

Facebook. Sus comentarios, aunque recibieron muchas burlas fuera de Silicon Valley, mandaron una clara señal dentro: no vamos a ceder.

Al cabo de unos días, Adam Mosseri —el exresponsable del canal de noticias, que ahora dirigía Instagram— dijo que, mientras que el 6 de enero «supondría un gran cambio» en la relación de la sociedad con la tecnología, era importante no caer en ultracorrecciones. «Toda nueva tecnología ha pasado por estas olas —dijo, trazando una serie de comparaciones que yo ya había oído en Silicon Valley—: Lo mismo ocurrió con el VHS. Lo mismo ocurrió con la escritura. Lo mismo ocurrió con las bicicletas.»⁹ Pero nadie supo decirme nunca qué genocidios se le atribuían al VHS.

Facebook y otras empresas pasaron de prometer que habían aprendido la lección y que iban a cambiar a insistir, incluso de una forma más estridente que antes del 6 de enero, en que todas las pruebas que apuntaban a su responsabilidad eran sencillamente erróneas. «Los sistemas de Facebook no están diseñados para recompensar el contenido provocativo»,¹⁰ escribió el responsable de relaciones públicas de Facebook, el ex viceprimer ministro británico Nick Clegg, en una publicación de blog de 5.000 palabras titulada «Tú y el algoritmo: bailar es cosa de dos». Según Clegg, en realidad eran los usuarios quienes favorecían el «contenido sensacionalista».

En gran medida, las compañías volvieron a las andadas. Las actuaciones contra la desinformación electoral disminuyeron de forma drástica a lo largo de 2021, como descubrió la organización de supervisión Common Cause, pues se permitió que las mentiras que socavaban la democracia «permanecieran en Facebook y Twitter hasta producir metástasis». ¹¹ Los movimientos nacidos en las redes sociales siguieron aumentando y fueron filtrándose en el tejido de la gobernación estadounidense.

A principios de 2022, en un estudio se halló que más de uno de cada nueve diputados estatales de todo Estados Unidos pertenecía al menos a un grupo de Facebook de extrema derecha. ¹² Muchos llevaban las conspiraciones e ideologías que habían surgido en internet a las leyes, con la aprobación de legislación que limitaba el derecho al voto, las políticas contra el covid y las protecciones del colectivo LGBT. En un momento de pánico alimentado por internet porque supuestamente había maestros que «introducían» a los alumnos en la homosexualidad, algunos aprobaron leyes que animaban a los niños a

grabar a sus profesores para obtener pruebas:¹³ una acción perturbadoramente parecida a la conspiración de YouTube sobre los «kits gais» que había provocado tanto revuelo en Brasil. El Partido Republicano de Texas, que controlaba el senado del estado, la cámara de representantes del estado y el cargo de gobernador, cambió su lema oficial por «Somos la tormenta», el grito de guerra de QAnon. En dos casos independientes, en Colorado y Míchigan, funcionarios electorales leales a QAnon fueron pillados manipulando los sistemas de votación.¹⁴ En las siguientes elecciones, en 2022, hubo candidatos vinculados con QAnon en las papeletas de veintiséis estados.¹⁵

Los gigantes de las redes sociales siguieron mostrando tal cantidad de mentiras y conspiraciones relacionadas con el covid que en julio Vivek Murthy, el director general de salud pública de Estados Unidos, publicó una advertencia formal en la que afirmaba que las plataformas «permitían que la desinformación intoxicase nuestro entorno informativo y asumían poca responsabilidad ante sus usuarios»,¹⁶ lo que él calificaba como una «crisis apremiante». Al día siguiente, Biden dijo que Facebook estaba «matando gente».¹⁷

Los demócratas, junto con otros muchos actores, depositaron buena parte de sus esperanzas en las investigaciones federales sobre regulación que se les haría a las compañías. La iniciativa había empezado, a finales de 2020, con una subcomisión antimonopolio de la Cámara de Representantes, bajo control demócrata, que había publicado un informe de 449 páginas ese octubre en el que recomendaba introducir una normativa de gran envergadura contra Facebook, Google, Amazon y Apple que en algunos casos fragmentaría a las empresas.¹⁸ Aunque el informe se centraba en las conductas monopolistas, los congresistas no escondían su preocupación por el control de las compañías sobre cuestiones que iban más allá del mercado, y alertaban: «Están en riesgo nuestra economía y nuestra democracia».

El Departamento de Justicia había presentado una demanda contra Google en octubre, en la que involucraba a la empresa y a YouTube en abusos relacionados como los que se mencionaban en el informe de la Cámara de Representantes. La Comisión Federal de Comercio había presentado una demanda parecida contra Facebook en diciembre, con lo que daba a entender que podría intentar fragmentar la compañía. Ambos casos, a los que se sumaron litigios paralelos

iniciados por varios fiscales generales de algunos estados, se prolongaron a lo largo de 2021 y 2022. Pero, aunque la aplicación de leyes antimonopolio puede ser una herramienta poderosa, es poco quirúrgica. Separar Facebook de Instagram, o Google de YouTube, debilitaría a las compañías, quizás incluso de forma drástica. Pero no cambiaría la naturaleza subyacente de sus productos. Ni eliminaría las fuerzas económicas e ideológicas que dieron lugar a esos productos.

Pese a la palabrería sobre la recuperación de la confianza, Silicon Valley demostró, en un episodio ocurrido tan solo un mes después de la investidura de Biden, que ejercería todo el peso de su poder contra sociedades enteras para disuadirlas de tomar medidas. Los organismos reguladores australianos habían actuado para atacar la mayor vulnerabilidad de Silicon Valley: los ingresos. En febrero de 2021, Facebook y Google tendrían que pagar a los medios de comunicación australianos por el derecho a mostrar enlaces a su trabajo.¹⁹ Las plataformas, al fin y al cabo, se apropiaban de los ingresos publicitarios del sector de la información comerciando con su periodismo. La nueva normativa contenía una cláusula poderosa. Si las compañías tecnológicas y las agencias de noticias no lograban acordar un precio antes del plazo marcado, unos árbitros públicos lo determinarían en su lugar. En la práctica, las normas favorecían a News Corp, el megaconglomerado dirigido por el australiano Rupert Murdoch, quien en 2016 había amenazado a Zuckerberg justo con una medida de ese tipo.

Sin embargo, con independencia de sus méritos, como muestra del poder que tenían los Gobiernos sobre las plataformas de redes sociales, los resultados fueron reveladores. Google, poco antes de agotarse el plazo, cerró un acuerdo con News Corp y otras empresas y cumplió la ley. Facebook se negó. En lugar de eso, una mañana, los australianos descubrieron que Facebook había bloqueado todo el contenido informativo. La fuente principal de información de todo un país —el 39 % de los australianos decían que se informaban por ese canal—²⁰ de repente no contenía noticias. También se habían apagado muchas otras cosas:²¹ los políticos que se presentaban a la reelección, los grupos que trabajaban con víctimas de violencia doméstica o el servicio meteorológico estatal. Incluso, en medio de una pandemia, los departamentos de sanidad del Gobierno. Al fin la compañía había hecho lo que no había hecho en Birmania durante un genocidio de

varios meses que, según acusaciones plausibles, habría facilitado. Ni en Sri Lanka ni en India. En ningún caso una espiral de violencia, por muy mortífera que fuera, había llevado a la compañía a apagar el interruptor, ni el de un solo componente de la plataforma. Pero la semana en que Australia amenazó sus ingresos, hubo apagón general.

Los australianos podían acceder, por supuesto, a las noticias o a las páginas web del Gobierno de forma directa. Aun así, Facebook, gracias a su diseño deliberado, se había hecho esencial y había enseñado a los usuarios a depender de su plataforma hasta el extremo de ser el objetivo final de las noticias y la información. Desaparecido el contenido informativo, el vacío lo llenaron los rumores y la desinformación. Evelyn Douek, una experta australiana que estudiaba la gobernanza de las plataformas de redes sociales en la Facultad de Derecho de Harvard, calificó el apagón como una acción «calculada para tener repercusiones y desmedida». ²²Human Rights Watch describió la intervención como «alarmante y peligrosa». ²³Un diputado australiano avisó de que bloquear el servicio meteorológico estatal podía afectar el acceso de los ciudadanos a unas actualizaciones que, en una semana en que se estaban produciendo inundaciones e incendios forestales, podían ser de vida o muerte. ²⁴Unos días después, el Gobierno australiano capituló y concedió a Facebook unas generosas excepciones a las nuevas normas.

Los Gobiernos europeos siguieron imponiendo multas y regulaciones. Los responsables políticos reconocían de forma implícita que no tenían mucha más capacidad para forzar cambios estructurales en Silicon Valley de la que tenían sus homólogos australianos o los políticos desafortunados que yo había conocido en Sri Lanka (muchos de los cuales, en 2021, ya no ocupaban sus cargos, pues habían sido sustituidos por un hombre fuerte ultranacionalista cuyo ascenso se debía, en parte, a que atizaba el odio y los rumores en internet). Pero los europeos tenían la esperanza de introducir políticas que pudieran convertirse en modelos para otras áreas del mundo. El Gobierno francés inauguró un centro para la regulación digital, para el que contrató, entre otros, a Guillaume Chaslot, el exingeniero de YouTube. Chaslot se ha dedicado sobre todo a encontrar formas de mostrar a los usuarios de internet cómo los algoritmos guían sus experiencias mientras navegan por las plataformas, incluso a desvelarles las elecciones y hábitos del sistema en tiempo real. Chaslot lo comparaba

con las etiquetas de los paquetes de cigarrillos que avisan a los consumidores de que fumar provoca cáncer de pulmón. A comienzos de 2022, la Unión Europea empezó a elaborar nuevas regulaciones que restringirían cómo las compañías tecnológicas estadounidenses podrían utilizar los datos personales de los consumidores europeos. Facebook, en un informe anual publicado en febrero, amenazó con una versión aún más dura de su estrategia en Australia advirtiendo de que, si la UE seguía adelante, «probablemente no podrían ofrecer en Europa muchos de nuestros productos y servicios más significativos, por ejemplo Facebook e Instagram».²⁵ El ministro de Finanzas alemán no se inmutó con la coerción, y dijo a los periodistas que, desde que había dejado las redes sociales cuatro años antes, «la vida ha sido fantástica». El ministro de Finanzas francés, casi retando a la compañía a cumplir sus amenazas, declaró: «Viviríamos muy bien sin Facebook».²⁶

La presión pública en Estados Unidos era más desigual. El expresidente Obama, en un discurso que dio en 2022 en el corazón de Silicon Valley, alertó de que las redes sociales estaban «propulsando algunos de los peores impulsos de la humanidad». Calificó «la profunda transformación [...] en la forma de comunicarnos y de consumir información» como uno de los principales factores de las adversidades, cada vez mayores, de la democracia, lo que anima a ciudadanos y Gobiernos a poner límites a las compañías. Aun así, la Administración Biden tenía todas las energías puestas en estabilizar la salud pública y en las emergencias económicas provocadas por la pandemia, y luego en la invasión rusa de Ucrania. Mientras tanto, las compañías contrataron a muchos de los analistas externos y académicos que les habían sacado los colores investigando sus productos. Lo habitual era que los fichasen aspirantes a reformadores que ocupaban cargos medios, quienes probablemente lo pensaban de verdad cuando decían que presionar a las compañías desde dentro era más factible para provocar cambios que hacerlo desde fuera. Y muchos cayeron a los brazos de Silicon Valley por culpa de un mercado de los empleos académicos en decadencia, donde las oportunidades para los doctores, cargados con deudas de estudios, se desintegraban del mismo modo en que los estadounidenses veían cómo los costes de la sanidad y la atención a la infancia se disparaban a raíz de la pandemia. Sean cuales fueran sus motivaciones para entrar

en Silicon Valley, como consecuencia de ello sus labores de divulgación pública cesaron.

Pero algunos, quizás gracias, en parte, a su notoriedad por los alarmantes descubrimientos que ya habían hecho sobre las redes sociales, consiguieron posiciones lo bastante estables en el mundo académico para seguir con su labor. El alemán Jonas Kaiser, el experto en los medios que había cartografiado YouTube, fue uno de ellos: empezó a trabajar en la facultad de la Universidad de Suffolk en Boston a finales de 2020. Otro era William Brady, que consiguió una plaza de profesor adjunto en la Universidad Northwestern en 2022. Ambos siguen tratando de entender las consecuencias de las redes sociales, lo mismo que decenas de personas como ellos, y aún les quedan descubrimientos por desvelar.

Algunos sí encontraron formas de presionar al sector desde dentro. Uno de ellos es Renée DiResta, que rastrea causas virtuales y actores malvados, como los antivacunas que motivaron el inicio de su andadura, en un centro de investigación de la Universidad de Stanford ubicado en el centro de Silicon Valley. Del mismo modo en que fue cauta, en su testimonio ante el Senado en 2018, e involucró a los demócratas que tenían mayor probabilidad de emprender acciones a la vez que querían encontrar puntos de coincidencia con los republicanos, DiResta ha aprovechado sus vínculos y los de Stanford con las empresas de redes sociales, lo que ha contribuido a organizar campañas exitosas para eliminar de las plataformas campañas de influencia vinculadas con Gobiernos, todo ello sin —hasta donde yo sé— renunciar ni un ápice a criticar públicamente esas mismas compañías.

Existe un grupo con la capacidad de influencia, el acceso y los conocimientos técnicos para presionar de manera eficaz a Silicon Valley: sus propios empleados. Después del 6 de enero, la indignación de los trabajadores no hizo más que aumentar: en marzo de 2021, un 40 % del personal de las grandes tecnológicas creía que gigantes como Google o Facebook deberían fragmentarse.²⁷ Pero, a pesar del músculo que exhibieron los trabajadores en 2020, su activismo se ha disipado en su mayor parte. La creación de sindicatos —los cinco dedos que se juntan en un puño— sigue siendo un anatema entre los empleados de Silicon Valley, lo que hace que organizarse en compañías de 50.000 trabajadores sea algo casi imposible. Y con unos salarios, unos

beneficios y una seguridad laboral casi inigualables en la época contemporánea, ¿por qué no ir a fichar durante algunos años, ganar una fortuna y jubilarse en paz?

Sin embargo, no todo el mundo eligió mirar hacia otro lado. En mayo de 2021, una empleada de Facebook llamada Frances Haugen, que trabajaba en un equipo que combatía la desinformación electoral en el extranjero, se hartó. Terminó pensando que sus superiores sacrificaban de forma deliberada la seguridad de los usuarios, de sociedades enteras, incluso de la democracia misma en aras de la maximización de los beneficios.

Haugen copió miles de informes de investigación internos, actas de reuniones y directivas empresariales fotografiando la pantalla en lugar de descargar los archivos directamente para evitar que se activasen medidas de seguridad. Durante meses, había estado hablando con Jeff Horwitz, un periodista del *Wall Street Journal* con quien se había puesto en contacto por primera vez en diciembre de 2020, impresionada por sus informaciones sobre la influencia nociva de las redes sociales en India, donde ella había vivido durante un tiempo.²⁸ Pero hasta pasados unos meses no decidió qué iba a hacer con los archivos.

«Sencillamente no quiero estar el resto de mi vida atormentándome por lo que no hice»,²⁹ dijo Haugen en un mensaje de texto a un amigo, en septiembre de 2021, cuando se planteaba si tenía que publicar la información. Poco después, mandó los archivos a la Comisión de Valores y Bolsa —un organismo regulador federal que supervisa las prácticas del mercado—, los cuales formarían parte de ocho querellas para denunciar conductas indebidas empresariales en Facebook. También los envió al Congreso y a Horwitz, quien empezó a hacer públicos los secretos que contenían.

En conjunto, los documentos contaban la historia de una compañía totalmente consciente de que sus perjuicios a veces superaban incluso las peores evaluaciones de los críticos. En algunos casos, los informes alertaban de forma explícita de peligros que tiempo después se volvieron mortíferos —como el fuerte aumento de los discursos de odio o de la desinformación sobre las vacunas— con suficiente antelación para que la compañía actuase y, si lo hacía (al final se negó), posiblemente salvase vidas. En informes inequívocos y con un lenguaje sin ambages, se veía que los propios datos de

Facebook y sus expertos confirmaban las acusaciones que la compañía había negado tan alegremente en público. A los ejecutivos de Facebook, entre ellos Zuckerberg, se les había dicho a las claras que su compañía representaba un peligro enorme, y esos ejecutivos habían intervenido una y otra vez para que sus plataformas siguieran operando a toda velocidad de todos modos. Los archivos, que Facebook minimizó alegando que no eran representativos, confirmaban en buena medida sospechas que se albergaban desde hacía tiempo. Pero algunos iban incluso más allá. En una presentación interna sobre cómo hacer que se engancharan más niños a los productos de Facebook se planteaba la siguiente pregunta: «¿Existe alguna forma de aprovechar las quedadas para jugar con el fin de impulsar la comunicación digital/crecimiento entre los niños?».

Mientras iba aumentando la indignación social, el programa *60 Minutes* anunció que emitiría una entrevista con la persona que había filtrado los documentos. Hasta ese momento, la identidad de Haugen se había mantenido en secreto. Su entrevista supuso un punto de inflexión en un debate, que entonces ya llevaba años en marcha, sobre esa tecnología por la claridad con que Haugen formulaba sus acusaciones: las plataformas amplificaban los perjuicios; Facebook lo sabía; la compañía tenía la capacidad de frenarlo pero elegía no hacerlo; y la compañía mintió continuamente a los organismos reguladores y a la sociedad. «Facebook se ha dado cuenta de que, si cambian el algoritmo para que sea más seguro —dijo Haugen—, la gente pasará menos tiempo en la plataforma, clicará menos anuncios y ellos ganarán menos dinero.»³⁰

Dos días después, Haugen testificó en la subcomisión del Senado. Se presentó como alguien que aspiraba a reformar el sector para salvaguardar su potencial. «Podemos tener unas redes sociales con las que disfrutemos y que nos conecten sin desgarrar nuestra sociedad, sin poner a nuestros hijos en peligro y sin sembrar violencia étnica por el planeta»,³¹ dijo a los senadores. En colaboración con Lawrence Lessig, experto de Harvard en asuntos jurídicos que le hacía de abogado de forma voluntaria, además de con una empresa de comunicación que Lessig contrató, Haugen también envió los documentos a dieciocho agencias de noticias estadounidenses y europeas. Informó a miembros del Congreso cuya presencia en la comisión les daba la potestad de elaborar nuevas regulaciones. Hizo una gira por capitales europeas,

donde se reunió con altos funcionarios cuyos Gobiernos habían sido pioneros en la regulación de las redes sociales.

Durante todo el proceso, Haugen hizo hincapié de forma sistemática en los errores de Facebook en países pobres. Ese historial, sostenía Haugen, ponía de manifiesto la insensibilidad de la compañía ante el bienestar de sus clientes, así como la capacidad desestabilizadora de la dinámica de la plataforma que, a fin de cuentas, tenía lugar en todas partes. «Lo que vemos en Birmania o en Etiopía —dijo en una comisión— no son sino los primeros capítulos de una novela que tiene un final mucho más siniestro que cualquier cosa que podamos leer.»³²

Los demócratas y las organizaciones de derechos humanos, aprovechando la indignación que Haugen generaba y la amenaza de los litigios antimonopolio que hay en marcha, se han centrado en hacer presión para que se apliquen regulaciones contundentes contra las empresas de redes sociales. En definitiva, el Gobierno estadounidense quizás sea el único ente que queda en el mundo con la capacidad suficiente para forzar a las compañías a cambiar. De todos modos, cuesta imaginar que varias decenas de diputados, que legislan en uno de los parlamentos más disfuncionales del mundo occidental, fueran capaces de seguir constantemente el ritmo de miles de ingenieros cuyo incentivo sería sustituir la función o la política que los congresistas hubieran sometido mediante una ley por algo que pudiera conseguir el mismo resultado.

Obligar a las compañías a regularse ellas mismas es también un camino incierto. Los gigantes de las redes sociales, tal como están constituidos en la actualidad, quizás simplemente sean incapaces de dar marcha atrás en sus peores tendencias. Desde un punto de vista técnico, sería sencillo. Pero las fuerzas culturales, ideológicas y económicas que llevaron a los ejecutivos a crear y sobrecargar esos sistemas siguen en pie. Políticos y activistas pueden gritar a Zuckerberg, Wojcicki y a los demás todo lo que quieran. Los creadores y directores ejecutivos de esas empresas, a pesar de su increíble riqueza, han sido —se hayan dado cuenta o no— prisioneros de sus creaciones desde el día en que un inversor de capital riesgo les extendió un cheque a cambio de la promesa de un crecimiento exponencial permanente.

Al preguntarle cuál sería la forma más eficaz de reformar tanto

las plataformas como las compañías que las controlan, Haugen dio una respuesta sencilla: desactivar el algoritmo. «Creo que no queremos que los ordenadores decidan a qué le prestamos atención», dijo. También sugirió que, si el Congreso recortaba las protecciones a la responsabilidad, lo que haría que las compañías fuesen responsables en términos legales de las consecuencias de todo lo que promovieran sus sistemas, «prescindirían de las clasificaciones basadas en la participación». Las plataformas volverían a la década del año 2000, cuando solo te mostraban las publicaciones de tus amigos por orden cronológico. Ninguna IA para abrumarte con un contenido que maximiza tu atención o que te precipita por madrigueras digitales.

Su respuesta seguía un patrón habitual que se ha ido vislumbrando en los años que me he pasado informando sobre las redes sociales. Cuanto más tiempo se dedica uno a estudiar las plataformas, sea cual sea su disciplina, más probable es que coincida con la respuesta de Haugen: hay que desactivar el algoritmo. A veces la recomendación es más restringida. Jonas Kaiser instó a YouTube a desactivar sus algoritmos en cuestiones sensibles como la salud o todo lo relacionado con los niños. A veces es más extensa. Benedict Evans, que fue inversor de capital riesgo en Andreessen Horowitz, propuso «eliminar capas enteras de aquellas mecánicas que posibilitan abusos». Al fin y al cabo, los algoritmos no son ni mucho menos la única característica que hay detrás del caos de las redes sociales. Unas interfaces al estilo de los casinos, los botones de compartición, los contadores públicos de «me gusta», las recomendaciones de grupos: todas estas cosas son intrínsecas a las plataformas y a sus perjuicios.

Hay también, como en cualquier aspecto que sea tan polémico como este, varios expertos que disienten y sostienen que se sobredimensionan los efectos de las redes sociales. No niegan las pruebas que constatan el papel de la tecnología en aspectos nocivos como la radicalización, pero emplean métodos diferentes que dan resultados más leves. Aun así, su visión es minoritaria y tiene una relevancia relativa, algo parecido a sostener que el papel de las emisiones de los vehículos en el cambio climático es menor que el de las centrales de carbón.

Casi ninguno de esos expertos o disidentes defiende que el mundo se beneficiaría de la eliminación total de las redes sociales. Al fin y al cabo, todos emprendieron sus estudios creyendo que las redes sociales

suponían un bien innegable y que, liberadas de su mecánica de fomento de los ingresos, aún podrían ser algo revolucionario. Pero, sea cual sea la recomendación, para una gran cantidad de investigadores serios, analistas o activistas por los derechos humanos, se trata de desactivar el algoritmo de alguna forma. Eso conduciría a un internet menos atractivo, menos cautivador, en el que los vídeos sorprendentes de YouTube o los grupos de Facebook emocionales serían menos frecuentes y costaría más acceder a ellos. Pero todas las pruebas disponibles sugieren también que sería un mundo con menos profesores que tendrían que esconderse huyendo de la persecución, menos familias quemadas vivas en sus casas debido a unos disturbios alimentados por rumores, menos vidas arruinadas por una infamia innecesaria o por la falsa promesa del extremismo. Menos niños privados de unas vacunas que les salvarían la vida o expuestos a una sexualización involuntaria. Tal vez incluso menos democracias destruidas por la polarización, las mentiras y la violencia.

Creo que un motivo por el que muchos expertos coinciden en esta respuesta es que muchos, cuando llevan cuatro o cinco grandes estudios, han ido a Silicon Valley a reunirse con los líderes supremos-ingenieros. Y se dan cuenta de que un aspecto de su defensa pública es cierto: creen que no están fomentando a propósito la desinformación, el odio o el tribalismo. Lo máximo que se plantean sobre esos efectos es la posibilidad de contenerlos. Pero eso es lo que hace que visitar Silicon Valley sea tan perturbador. Hay una mezcla de ideología, avaricia y la opacidad tecnológica de un complejo aprendizaje automático que ciega a los ejecutivos y les impide ver sus creaciones en su totalidad. En los aspectos que importan, las máquinas, en esencia, están desgobernadas.

Cuanto más hablaba con psicólogos y analistas de redes, reguladores e ingenieros reformados, más me recordaban los términos que empleaban para describir esta tecnología a HAL 9000, la inteligencia artificial de *2001: Una odisea del espacio*, la película de Kubrick que a lo largo de mis investigaciones sobre las redes sociales me había ido pareciendo relevante por sus paralelismos con la actualidad. En el filme, HAL, aun siendo responsable de la seguridad de la tripulación, sobreinterpreta el código informático que le ordena asegurar la llegada a la destinación prevista cueste lo que cueste e intenta matar a todas las personas que hay a bordo. HAL no está

pensada como un villano. Si hubiera alguno, serían los ingenieros que, con su soberbia, supusieron que las acciones de su creación serían tan benévolas como sus intenciones, o quizás los astronautas que entregaron su confianza a una máquina que puede decidir entre la vida o la muerte, con unos incentivos que podían no coincidir con los suyos.

La lección de *2001: Una odisea del espacio* no era ni mucho menos que hubiera que actualizar a HAL con más modificaciones algorítmicas con la esperanza de que, la próxima vez, pudiera comportarse de una forma un tanto más responsable. Ni tampoco que los ingenieros de HAL debieran pedir perdón y prometer que lo harían mejor. Y, desde luego, no era que el fabricante de HAL asumiera un control cada vez mayor de las vidas de sus clientes mientras políticos y periodistas reflexionasen sobre la naturaleza del robot. La lección era inequívoca: desactivad a HAL. Aunque eso significase perder los beneficios que aportase. Aunque fuera difícil —como se ve en las escenas finales de la película— arrancar los tentáculos de HAL de los sistemas que gobiernan todas las facetas de la vida de los astronautas. Aunque la máquina se defendiese con todas sus fuerzas.

Agradecimientos

Todo valor que pueda tener este libro se debe en gran medida a su editor, Ben George, cuya pasión, esmero y espíritu contagioso no flaquearon ni un solo instante, ni siquiera durante maratónicas sesiones de edición de dieciséis horas con un bebé en el regazo. Gracias también a Bruce Nichols, Katharine Myers y a todas las personas de la editorial Little, Brown por su entusiasmo y su fe.

Este libro no habría sido posible sin Jennifer Joel, de ICM Partners, quien lo condujo por entre oportunidades y obstáculos con tenacidad y buen juicio.

Estoy en deuda con muchos de mis compañeros del *New York Times*. Michael Slackman, jefe de la sección Internacional, fue el primero en sugerir la idea que se convertiría en este proyecto y, además, supervisó y apoyó actividades periodísticas que darían lugar a apartados de este libro o servirían de inspiración para ellos. Eric Nagourney, Juliana Barbassa y Doug Shorzman editaron esos trabajos echando sangre y sudor a unos textos que de hecho deberían llevar su nombre. Amanda Taub, con quien empecé la columna «The Interpreter» en 2016, ha sido coautora de reportajes sobre Sri Lanka, Alemania y Brasil. También tuve la suerte de trabajar o compartir autoría periodística con Wai Moe en Birmania; Katrin Bennhold y Shane Thomas McMillan en Alemania; Dharisha Bastians en Sri Lanka; y Mariana Simões y Kate Steiker-Ginzberg en Brasil; junto con Alyse Shorland y Singeli Agnew como productoras de *The Weekly*. Pui-Wing Tam, Kevin Roose, Paul Mozur, entre otros, me brindaron su apoyo y solidaridad al informar sobre las redes sociales. Gracias a los directores del *New York Times* por apoyar esas labores periodísticas y por concederme el espacio para escribir este libro.

Un sinnúmero de académicos, investigadores y otras personas dieron gratuitamente su energía, ideas y a veces trabajos originales para este proyecto. Como lo hicieron también médicos, ingenieros, activistas por los derechos humanos y otras personas que están en las trincheras

de las cuestiones tratadas en el libro, además de las personas a las que de un modo eufemístico llamamos «fuentes»: supervivientes, filtradores, testigos. En muchos sentidos este libro es suyo.

Gracias, en especial, a mi madre, a mi padre y a Joanna por creer en mí y por convertirme en la persona que soy. Y a Jordan por no dejarme desfallecer, ni física ni psicológicamente, y por hacer que todo esto merezca la pena.

Notas

1. «Facebook Just Opened an Epic \$300 Million Gehry-Designed Campus with a Redwood Forest and Rooftop Garden», Andrew Evers, CNBC, 4 de septiembre de 2018.

2. «Tina Vaz on Facebook's Artist in Residence Program», Whitewall.art, 15 de enero de 2020.

3. Entrevista con Monika Bickert, directora de gestión de políticas internacionales de Facebook, en octubre de 2018.

4. Entrevista con Nathaniel Gleicher, director de políticas de seguridad de Facebook, en octubre de 2018.

5. Entrevista con Sara Su, entonces directora de gestión de producto sobre el equipo responsable del canal de noticias, en octubre de 2018.

6. Extraído de «Facebook Executives Shut Down Efforts to Make the Site Less Divisive», Jeff Horwitz y Deepa Seetharaman, *The Wall Street Journal*, mayo de 2020.

7. «Facebook's Civil Rights Audit — Final Report», Laura W. Murphy y el bufete de abogados Relman Colfax, <https://www.facebook.com/about/civilrights>, 8 de julio de 2020.

1. «How Facebook and YouTube Help Spread Anti-Vaxxer Propaganda», Julia Carrie Wong, *The Guardian*, 1 de febrero de 2019.

2. *The Code: Silicon Valley and the Remaking of America*, Margaret O'Mara, 2019, expone con gran detalle cómo los rasgos y personalidades fundacionales de Silicon Valley crearon las redes sociales actuales y las convirtieron en lo que son.

3. Extraído de *The Man Behind the Microchip: Robert Noyce and the Invention of Silicon Valley*, Leslie Berlin, 2005, una de las muchas biografías de los contemporáneos de Shockley en la que se detallan sus insultos. Para más detalles, en concreto sobre su giro hacia la eugenesia y el racismo, véase *Broken Genius: The Rise and Fall of William Shockley, Creator of the Electronic Age*, Joel N. Shurkin, 2006.

4. O'Mara, 2019, pp. 7-9.

5. Entrevista con Leslie Berlin, historiadora de la Universidad de Stanford University, en mayo de 2020.

6. Fuente de la base de usuarios de Facebook: «Inside Mark Zuckerberg's Controversial Decision to Turn Down Yahoo's \$1 Billion Early Offer to Buy Facebook», Mike Hoefflinger, *Business Insider*, 16 de abril de 2017. De Friendster: «The Friendster Autopsy: How a Social Network Dies», Robert McMillan, *Wired*, 27 de febrero de 2013, y «Friendster Patents Social Networking», Pete Cashmore, *Mashable*, 7 de julio de 2006. De Orkut: «Google's Orkut Captivates Brazilians», Seth Kugel, *The New York Times*, 9 de abril de 2006. De Myspace: «The Decline of Myspace: Future of Social Media», Karl Kangur, *DreamGrow*, 13 de agosto de 2012.

7. *Facebook: The Inside Story*, Steven Levy, 2020, presenta informaciones detalladas y de primera mano de Zuckerberg y otros empleados de alto rango sobre la decisión de rechazar la oferta de Yahoo, además del episodio posterior del canal de noticias.

8. Quizás donde mejor se retrata este fenómeno, que los científicos conocen como conocimiento general, es en «How Does Media Influence Social Norms? Experimental Evidence on the Role of Common Knowledge», Eric Arias, *Political Science Research and Methods*, julio de 2019. Véase también *Rational Ritual: Culture, Coordination, and Common Knowledge*, Michael Suk-Young Chwe, reedición de 2013, o el trabajo de Betsy Levy Paluck, de la Universidad de Princeton, que se explora más adelante.

9. «Calm down. Breathe. We hear you», Mark Zuckerberg sobre las ya desaparecidas Facebook Notes, septiembre de 2006.

10. «Microsoft Buys Stake in Facebook», Brad Stone, *The New York Times*, 25 de octubre de 2007.

11. Todas las estadísticas que aparecen en este párrafo proceden de *Social Media Fact Sheet*, Pew Research Center, 7 de abril de 2021.

12. Las estadísticas sobre el uso de Facebook proceden de «10 Facts About Americans and Facebook», John Gramlich, Pew Research Center, 1 de junio de 2021. Las estadísticas sobre la socialización en persona son del Bureau of Labor Statistics y de «Facebook Has 50 Minutes of Your Time Each Day. It Wants More», James B. Stewart, *The New York Times*, 5 de mayo de 2016.

13. Los comentarios de Parker proceden de un congreso empresarial organizado por el portal Axios en noviembre de 2017, donde fue entrevistado por el periodista Mike Allen.

14. «The Formula for Phone Addiction Might Double as a Cure», Simone Stolzoff, *Wired*, 1 de febrero de 2018.

15. La descripción de la dopamina y sus usos y abusos, así como las referencias a Pavlov, proceden del artículo académico «A Selective Role for Dopamine in Stimulus-Reward Learning», Shelly B. Flagel *et al.*, *Nature*, 2011.

16. Para tener una visión general asequible de los descubrimientos de Skinner, así como más información sobre los amplificadores duales de las recompensas variables e intermitentes (las redes sociales proporcionan ambos elementos), véase «Schedules of Reinforcement», Annabelle G.Y. Lim, *Simply Psychology*, 2 de julio de 2020. Pueden encontrarse más detalles, junto con citas que fundamentan las investigaciones neurológicas, en el capítulo 2 de *Behave: The Biology of Humans at Our Best and Worst*, Robert M. Sapolsky, 2017.

17. «The Top 10 Valuable Facebook Statistics», Zephoria Research, 2021.

18. «47 Facebook Stats That Matter to Marketers in 2021», Christina Newberry, *Hootsuite*, 11 de enero de 2021.

19. «The Welfare Effects of Social Media», Hunt Allcott, Luca Braghieri, Sarah Eichmeyer y Matthew Gentzkow, *American Economic Review*, marzo de 2020.

20. «What's the History of the 'Awesome Button' (That Eventually Became the Like Button) on Facebook?», Andrew Bosworth, *Quora*, 16 de octubre de 2014.

21. «Sociometer Theory», de Mark R. Leary, es el capítulo 33 de *Handbook of Theories of Social Psychology*, vol. 2, 2011.

22. Más adelante en este libro encontraréis una explicación más completa de esta cifra, sus orígenes e implicaciones. Su progenitor, Robin Dunbar, hace un útil resumen en «Dunbar's Number: Why My Theory That Humans Can Only Maintain 150 Friendships Has Withstood 30 Years of Scrutiny», *The Conversation*, 12 de mayo de 2021.

23. *Survival of the Friendliest*, Brian Hare y Vanessa Woods, 2021.

24. Puede encontrarse una explicación más detallada de los orígenes y funciones exclusivamente humanos de estas emociones en *Humankind: A Hopeful History*, Rutger Bregman, 2019.

25. «The Binge Breaker», Bianca Bosker, *The Atlantic*, noviembre de 2016.

26. «Jack Dorsey on Twitter's Mistakes», The Daily, pódcast del *New York Times*, 7 de agosto de 2020.

27. A menos que se indique lo contrario, todas las referencias a los efectos neurológicos del uso de las redes sociales en esta sección se basan en las investigaciones de Dar Meshi, neurocientífico de la Universidad Estatal de Michigan. Véase, en concreto, «The Emerging Neuroscience of Social Media», Meshi *et al.*, en *Trends in Cognitive Sciences*, diciembre de 2015.

28. «Facebook Usage on Smartphones and Gray Matter Volume of the Nucleus Accumbens», Christian Montag *et al.*, *Behavioural Brain Research*, junio de 2017.

29. «The Inventor of the ‘Like’ Button Wants You to Stop Worrying About Likes», Julian Morgans, *Vice News*, 5 de julio de 2017.

30. «‘Our Minds Can Be Hijacked’: The Tech Insiders Who Fear a Smartphone Dystopia», Paul Lewis, *The Guardian*, 6 de octubre de 2017.

31. «Individuals and Groups in Social Psychology», Henri Tajfel, *British Journal of Social & Clinical Psychology* 18, n.º 2, 1979.

32. Puede encontrarse un buen resumen en «The Social Identity Theory of Intergroup Behavior», Henri Tajfel y John C. Turner, *Psychology of Intergroup Relations*, 1986.

33. «Social Psychology of Intergroup Relations», Henri Tajfel, *Annual Review of Psychology* 33, 1982.

34. «Social Categorization and Discriminatory Behavior: Extinguishing the Minimal Intergroup Discrimination Effect», Anne Locksley, Vilma Ortiz y Christine Hepburn, *Journal of Personality and Social Psychology* 39, n.º 5, 1980.

35. Puede encontrarse un resumen de muchos estudios de este tipo en «Ingroup Favoritism and Prejudice», en *Principles of Social Psychology*, First International Edition, Charles Stangor, Hammond Tarry y Rajiv Jhangiani, 2014.

36. Heston ha contado esto en varias entrevistas, la primera de las cuales se publicó en «The Arts», Jeff Rovin, *Omni Magazine*, noviembre de 1980, p. 140. Natalie Trundy, coprotagonista de la película, describió por separado el incidente al menos una vez: «The Day of the Apes», Tom Weaver, revista *Starlog*, septiembre de 2001, p. 20.

37. Tajfel, «Social Psychology of Intergroup Relations».

38. Este efecto se ha demostrado en repetidas ocasiones, quizás la mayor demostración fuera en «A New Stress-Based Model of Political Extremism», Daphna Canetti-Nisim, Eran Halperin, Keren Sharvit y Stevan E. Hobfoll, *Journal of Conflict Resolution* 53, n.º 2, junio de 2009.

39. *Why We're Polarized*, Ezra Klein, 2020, p. 143.

40. «Google Chief Says Internet Freedom Key for Myanmar», vídeo de Agence France-Presse, 22 de marzo de 2013.

41. Mencionado en «Fears over Facebook Regulation Proposal», Tim McLaughlin, *Myanmar Times*, 15-21 de julio de 2013.

42. *Sticks and Stones: Hate Speech Narratives and Facilitators in Myanmar*, del grupo de investigación sin ánimo de lucro C4ADS, febrero de 2016.

43. *Internet Health Report: 2019*, publicado por la Mozilla Foundation, ofrece datos de encuestas que muestran que la mayoría de los usuarios de varios países con programas de datos gratuitos «no tienen ni idea de que hay un internet más allá de Facebook».

44. Véase el informe que aparece en *Sticks and Stones*, de C4ADS, para encontrar detalles profundos y a menudo perturbadores sobre los discursos de odio que circulaban por Facebook en Birmania en esa época.

45. «How Facebook's Rise Fueled Chaos and Confusion in Myanmar», Timothy McLaughlin, *Wired*, julio de 2018, describe este episodio con detalle. Véase también «Why Facebook Is Losing the War on Hate Speech in Myanmar», Steve Stecklow, Reuters, 15 de agosto de 2018.

1. *Crash Override: How Gamergate (Nearly) Destroyed My Life, and How We Can Win the Fight Against Online Hate*, Zoë Quinn, 2017, p. 2.

2. Se puede acceder a registros de estos debates en «GamerGate — #GameOverGate IRC Logs Explanation», Knowyourmeme.com, sin fecha. Véase también «Zoe Quinn's Screenshots of 4chan's Dirty Tricks Were Just the Appetizer. Here's the First Course of the Dinner, Directly from the IRC log», David Futrelle, Wehuntedthetmammoth.com, 8 de septiembre de 2014.

4. «Game of Fear», Zachary Jason, *Boston Magazine*, 28 de abril de 2015.

5. «Zoë and the Trolls», Noreen Malone, *The New York Magazine*, julio de 2017.

6. «That Time the Internet Sent a SWAT Team to My Mom's House», Caroline Sindors, Narratively.com, 17 de julio de 2015.

7. «His 'Swatting' Call Led to the Death of a Man. Now He is Going to Prison for 20 Years», Steve Almasy y Melissa Alonso, CNN, 30 de marzo de 2019.

8. «FBI Arrests Man Suspected of Orchestrating Dozens of ‘Swatting’ Calls», Timothy B. Lee, Arstechnica.com, 14 de enero de 2020.

9. «Intel Pulls Ads from Gamasutra, and Then Apologizes for It», Dean Takahashi, Venturebeat.com, 3 de octubre de 2014.

10. «I'm Brianna Wu, and I'm Risking My Life Standing Up to Gamergate», Brianna Wu, *Daily Dot*, 12 de febrero de 2015.

11. O'Mara, pp. 90-92.

12. Berlin, p. 194.

14. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*, Fred Turner, 2010, pp. 71-72.

15. Cliff Figallo, uno de los arquitectos de la plataforma, ha dicho, por ejemplo, que «los principios de tolerancia e inclusión, una asignación justa de los recursos, la responsabilidad distribuida, una gestión basada en el ejemplo y la influencia, una jerarquía organizativa horizontal, la formulación colaborativa de las políticas y la aceptación de unos valores libertaristas cercanos al anarquismo eran remanentes de nuestra experiencia de vida comunitaria». Fuente: Turner, p. 148.

17. «A Cloudy Crystal Ball / Apocalypse Now», presentación de David Clark, julio de 1992, en el 24º congreso anual de la Internet Engineering Task Force.

18. «A Declaration of the Independence of Cyberspace», John Perry Barlow, 8 de febrero de 1996. Al principio se difundió a montones de páginas web de forma simultánea, ahora está disponible en <<https://www.eff.org/cyberspace-independence>>.

19. «Twitter's Tony Wang: 'We Are the Free Speech Wing of the Free Speech Party'», Josh Halliday, *The Guardian*, 22 de marzo de 2012.

20. Levy, *Facebook: The Inside Story*, p. 458.

21. Esto es así según Dave Morin, exingeniero de Facebook, parafraseado en Levy, p. 149.

22. «The Facebook Dilemma», *PBS Frontline*, 29 de octubre de 2018.

23. «Zuckerberg's Letter to Investors», Reuters, 1 de febrero de 2012.

25. *Hackers and Painters*, Paul Graham, 2004, p. 9.

26. «What We Look for in Founders», Paul Graham, Paulgraham.com, octubre de 2010.

27. «What I Did This Summer», Paul Graham, Paulgraham.com, octubre de 2005.

28. *Zero to One: Notes on Startups, or How to Build the Future*, Peter Thiel y Blake Masters, 2014, p. 40.

30. Pueden encontrarse pantallazos que documentan el incidente en «Kenny Glenn Case / Dusty the Cat», Knowyourmeme.com, 10 de septiembre de 2011.

31. «Media Manipulation and Disinformation Online», Alice Marwick y Rebecca Lewis, *Data & Society*, mayo de 2017.

32. En español se usa un término parecido, *nini*, en referencia a alguien que ni estudia ni trabaja. (*N. del t.*)

33. «The Trolls Among Us», Mattathias Schwartz, *The New York Times Magazine*, 3 de agosto de 2008.

34. *Behave: The Biology of Humans at Our Best and Worst*, Robert M. Sapolsky, 2017, pp. 163-164.

35. *Teens, Social Media, and Privacy*, Mary Madden *et al.*, Pew Research Center, 21 de mayo de 2013.

36. «From LOL to LULZ, the Evolution of the Internet Troll over 24 Years», Kristen V. Brown, Splinternews.com, 18 de marzo de 2016.

37. «How the Internet Beat Up an 11-Year-Old Girl», Adrian Chen, *Gawker*, 16 de julio de 2010.

38. «Myspace-Famous Musician Dahvie Vanity Was Accused of Child Sex Abuse for Years. Now the FBI is Involved», Kat Tenbarger, Insider.com, 2 de julio de 2020.

40. *You Are Here: A Field Guide for Navigating Polarized Speech, Conspiracy Theories, and Our Polluted Media Landscape*, Whitney Phillips y Ryan M. Milner, 2021, p. 58.

44. «The End of Kindness: Weev and the Cult of the Angry Young Man», Greg Sandoval, *The Verge*, 12 de septiembre de 2013.

45. «We're Awarding Goatse Security a Crunchie Award for Public Service», Michael Arrington, *TechCrunch*, 14 de junio de 2010.

46. «Lulz and Leg Irons: In the Courtroom with Weev», Molly Crabapple, *Vice News*, 19 de marzo de 2013.

47. «The Kleiner Perkins sFund: A \$250 Million Bet That Social Is Just Getting Started», Michael Arrington, *TechCrunch*, 21 de octubre de 2010.

48. «CEO 2.0», discurso de Bing Gordon en la cumbre Endeavor Entrepreneur Summit celebrada en San Francisco, California, el 28 de junio de 2011.

49. Puede encontrarse un relato exhaustivo de esta historia en «No Girls Allowed», Tracey Lien, Polygon.com, 2 de diciembre de 2013.

50. «How the Alt-Right's Sexism Lures Men into White Supremacy», Aja Romano, Vox, 26 de abril de 2018.

51. [Facebook: The Inside Story](#), Steven Levy, 2020, p. 213.

52. «Coevolution of Neocortical Size, Group Size, and Language in Humans», Robin Dunbar, *Behavioral and Brain Sciences* 16, 1993.

53. Esto es lo que decía la página de estadísticas de Facebook, ya desaparecida, en 2010. Véase, por ejemplo, «10 Fascinating Facebook Facts», *Mashable*, 22 de julio de 2010.

54. «Friends, Friendsters, and Top 8: Writing Community into Being on Social Network Sites», Danah Boyd, *First Monday* 11, n.º 12, diciembre de 2006.

55. «Like, How Many Friends Does Facebook Need?», Edo Elan, *The Product Guy*, 10 de mayo de 2010.

56. Zuckerberg había dicho: «Hay ese famoso número de Dunbar: que los seres humanos tienen la capacidad de mantener relaciones empáticas con unas ciento cincuenta personas. Yo creo que Facebook va a ampliar esa cifra». Extraído de Levy, *Facebook: The Inside Story*, p. 226.

1. «New Survey Reflects Lack of Women and Minorities in Senior Investment Roles at Venture Capital Firms», comunicado de prensa de la National Venture Capital Association, 14 de diciembre de 2016.

2. «Funding for Female Founders Stalled at 2.2% of VC Dollars in 2018», Emma Hinchliffe, *Fortune*, 28 de enero de 2019.

3. Literalmente, ‘cebo de cárcel’. Es un término usado para referirse a chicas menores de la edad de consentimiento, con quienes mantener relaciones sexuales tendría por tanto consecuencias penales. (*N. del t.*)

4. «Reddit CEO Addresses Violentacruz Controversy», Sean Hollister, *The Verge*, 19 de octubre de 2012.

5. Pueden encontrarse secciones de la publicación, que Reddit eliminó, en «Reddit's Confusing Response to the Distribution of Nudes», Alex Goldman, NPR, 8 de septiembre de 2014.

6. «From 1 to 9,000 Communities, Now Taking Steps to Grow Reddit to 90,000 Communities (and Beyond!)», Ellen Pao et al., Reddit, 24 de febrero de 2015.

7. «How Reddit Became a Worse Black Hole of Violent Racism than Stormfront», Keegan Hanks, *Gawker*, 10 de marzo de 2015.

8. «Reddit's Racists 'Celebrate' Charleston Terror — and Worry About the Blowback», Jacob Siegel, *The Daily Beast*, 12 de julio de 2017.

9. «Weev and the Rise of the Nazi Troll Army», Andrew Anglin, *Daily Stormer*, 4 de octubre de 2014.

10. Esta es la opinión de Brad Griffin, activista de extrema derecha. «Dylann Roof, 4chan, and the New Online Racism», Jacob Siegel, *Daily Beast*, 14 de abril de 2017.

11. *A Comparative Study of White Nationalist and ISIS Online Social Media Networks*, J. M. Berger, Programa sobre extremismo de la Universidad George Washington, septiembre de 2016.

12. «The Trolls Are Winning the Battle for the Internet», Ellen Pao, *The Washington Post*, 16 de julio de 2015.

13. «How Stephen Bannon Made Milo Dangerous», Keegan Hanks, Southern Poverty Law Center, 23 de febrero de 2017.

15. *Devil's Bargain: Steve Bannon, Donald Trump, and the Nationalist Uprising*, Joshua Green, 2017, p. 147.

16. «Here's How Breitbart and Milo Smuggled White Nationalism into the Mainstream», Joseph Bernstein, *BuzzFeed News*, 5 de octubre de 2017.

17. «An Establishment Conservative's Guide to the Alt-Right», Allum Bokhari y Milo Yiannopoulos, *Breitbart*, 29 de marzo de 2016.

18. «Behind the Racist Hashtag That Is Blowing Up Twitter», Joseph Bernstein, *BuzzFeed News*, 27 de julio de 2015.

19. «Meme Magic: Donald Trump Is the Internet's Revenge on Lazy Elites», Milo Yiannopoulos, *Breitbart*, 4 de mayo de 2016.

20. «This Is The Daily Stormer's Playbook», Ashley Feinberg, *HuffPost*, 13 de diciembre de 2017.

22. *Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election*, Robert M. Faris *et al.*, artículo de investigación del Berkman Klein Center for Internet & Society, 2017.

23. «Key Findings About the Online News Landscape in America», A. W. Geiger, *Pew Research Center*, 11 de septiembre de 2019.

24. «Mark Zuckerberg Is Struggling to Explain Why Breitbart Belongs on Facebook News», Adi Robertson, *The Verge*, 25 de octubre de 2019.

25. Todos los resultados que aparecen en este párrafo y el siguiente proceden de Faris *et al.*

1. «Full transcript: Walter Palmer Speaks About Cecil the Lion Controversy», Paul Walsh, *Minneapolis Star Tribune*, 7 de septiembre de 2015.

3. «Zimbabwe's 'Iconic' Lion Cecil Killed by Hunter», BBC News, 27 de julio de 2015.

4. «Odeo Releases Twttr», Michael Arrington, *TechCrunch*, 15 de julio de 2006.

5. «twtr sketch», Jack Dorsey, Flickr, 24 de marzo de 2006.

6. En lo que fue una reacción representativa, Mark Pfeifle, un asesor adjunto de seguridad nacional, de la Casa Blanca de George W. Bush, instó a que se concediese el Premio Nobel de la Paz a Twitter, plataforma sin la cual, escribió Pfeifle, «el pueblo iraní no se habría sentido capacitado y seguro de sí mismo para luchar por la libertad y la democracia». «A Nobel Peace Prize for Twitter?», Mark Pfeifle, *Christian Science Monitor*, 6 de julio de 2009.

7. «Outrage and Backlash: #CecilTheLion Racks Up 670K Tweets in 24 Hours», Jordan Valinsky, Digiday.com, 29 de julio de 2015.

8. «The Entire World Is Enraged with Walter Palmer, the American Dentist Who Killed Cecil the Lion», Hanna Kozłowska, QZ.com, 28 de julio de 2015.

9. «Stars Blast Minnesota Dentist over Killing of Cecil the Lion», Justin Ray, NBC News, 31 de julio de 2015.

10. «Killer of Cecil the Lion Finds Out That He Is a Target Now, of Internet Vigilantism», Christina Capecchi y Katie Rogers, *The New York Times*, 29 de julio de 2015.

11. Comentario sobre «Meanwhile, Outside Walter Palmer's Dentistry Office» escrito por el usuario CinnamonDolceLatte, Reddit, 29 de julio de 2015.

12. «Where Clicks Reign, Audience Is King», Ravi Somaiya, *The New York Times*, 16 de agosto de 2015.

13. «The Clickbait Candidate», James Williams, Quillette, 3 de octubre de 2016.

15. «The New Synthesis in Moral Psychology», Jonathan Haidt, *Science*, 18 de mayo de 2007.

16. «How Infants and Toddlers React to Antisocial Others», Kiley Hamlin *et al.*, *Proceedings of the National Academy of Sciences*, 13 de diciembre de 2011. Véase también *Just Babies: The Origins of Good and Evil* (2013), de Paul Bloom, coautor de los estudios de Hamlin.

18. «The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment», Jonathan Haidt, *Psychological Review*, octubre de 2001.

19. «A Dongle Joke That Spiraled Way Out of Control», Kim Mai-Cutler, *TechCrunch*, 21 de marzo de 2013.

20. *Dongle* (de ahí, *Donglegate*) puede traducirse como ‘adaptador’ —un dispositivo que se conecta a otro—, pero también puede significar ‘pene’ en lenguaje coloquial. (*N. del t.*)

21. «Re-Shaming the Debate: Social Norms, Shame, and Regulation in an Internet Age», Kate Klonick, *Maryland Law Review* 76, n.º 4, 2016.

23. «How One Stupid Tweet Blew Up Justine Sacco's Life», Jon Ronson, *The New York Times Magazine*, 12 de febrero de 2015.

24. «The Nazi Salute Picture That Divided an American Town», Chris McGreal, *The Guardian*, 10 de enero de 2019.

25. «Twitter Hates Me. The Des Moines Register Fired Me. Here's What Really Happened», Aaron Calvin, *Columbia Journalism Review*, 4 de noviembre de 2019.

26. «The CEO of Holy Land Hummus I Know Doesn't Match the Social Media Monster», Rob Eshman, *The Forward*, 8 de junio de 2020.

27. «Bogus Social Media Outrage Is Making Authors Change Lines in Their Books Now», Laura Miller, *Slate*, 8 de junio de 2021.

28. «Inside a Battle Over Race, Class, and Power at Smith College», Michael Powell, *The New York Times*, 24 de febrero de 2021.

29. «Anger Is More Influential than Joy: Sentiment Correlation in Weibo», Rui Fan *et al.*, *PLOS One* 9, n.º 10, octubre de 2014.

30. «Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks», Adam D.I. Kramer *et al.*, *Proceedings of the National Academy of Sciences* 111, n.º 24, junio de 2014. «Hostile Emotions in News Comments: A Cross-National Analysis of Facebook Discussions», Edda Humprecht *et al.*, *Social Media + Society* 6, n.º 1, marzo de 2020. «Behavioral Effects of Framing on Social Media Users: How Conflict, Economic, Human Interest, and Morality Frames Drive News Sharing», Sebastián Valenzuela *et al.*, *Journal of Communication* 67, n.º 5, octubre de 2017. «Emotion Shapes the Diffusion of Moralized Content in Social Networks», William J. Brady *et al.*, *Proceedings of the National Academy of Sciences* 114, n.º 28, julio de 2017. «Critical Posts Get More Likes, Comments, and Shares than Other Posts», *Pew Research Center*, 21 de febrero de 2017.

31. Puede encontrarse un resumen del experimento y sus resultados en «How to Tame a Fox and Build a Dog», Lee Alan Dugatkin y Lyudmila Trut, *American Scientist*, julio-agosto de 2017. «The Silver Fox Domestication Experiment», Lee Alan Dugatkin, *Evolution: Education and Outreach* 11, 2018.

32. «Early Canid Domestication: The Farm-Fox Experiment», Lyudmila Trut, *American Scientist* 87, n.º 2, marzo-abril de 1999.

33. La extrapolación de las investigaciones de Trut sobre esta cuestión y las lecciones inmediatamente posteriores para la antropología humana y la conducta son la base del libro de Richard Wrangham *The Goodness Paradox: The Strange Relationship Between Virtue and Violence in Human Evolution*, 2019. Aunque estas conexiones y conclusiones no son ni mucho menos únicas de Wrangham, él es la persona que se asocia más a menudo con la teoría global.

35. *Conditions of Liberty: Civil Society and Its Rivals*, Ernest Gellner, 1994.

37. «Evolution of Coalitionary Killing», Richard Wrangham, *Yearbook of Physical Anthropology* 42, n.º 1, 1999. Véase también Wrangham, *Goodness Paradox*, p. 244.

38. «Perceptual Dehumanization of Faces Is Activated by Norm Violations and Facilitates Norm Enforcement», Katrina M. Fincher y Philip E. Tetlock, *Journal of Experimental Psychology* 145, n.º 2, 2016.

39. «Moral Grandstanding: There's a Lot of It About, All of It Bad», Justin Tosi y Brandon Warmke, *Aeon*, 10 de mayo de 2017.

41. «Moral Grandstanding in Public Discourse: Status-Seeking Motives as a Potential Explanatory Mechanism in Predicting Conflict», Joshua B. Grubbs *et al.*, *PLOS One* 14, n.º 10, 2019.

42. «Reputation Fuels Moralistic Punishment That People Judge To Be Questionably Merited», Jillian J. Jordan y Nour S. Kteily, documento de trabajo, 2020. Véase también «Signaling When No One Is Watching: A Reputation Heuristics Account of Outrage and Punishment in One-Shot Anonymous Interactions», Jillian J. Jordan y D. G. Rand, *Journal of Personality and Social Psychology* 118, n.º 1, 2020.

43. Puede encontrarse un relato completo del incidente en «How Two Lives Collided in Central Park, Rattling the Nation», Sarah Maslin Nir, *The New York Times*, 14 de junio de 2020.

44. «The Bird Watcher, That Incident and His Feelings on the Woman's Fate», Sarah Maslin Nir, *The New York Times*, 27 de mayo de 2020.

45. De las muchas reflexiones sobre el cambio en las costumbres sociales impulsadas por las redes sociales, esta es quizás de las que mejor captan la ambivalencia de ese momento: «Karens All the Way Down», Kat Rosenfield, *Arc Digital*, 26 de mayo de 2020.

46. «Political Rumoring on Twitter During the 2012 US Presidential Election: Rumor Diffusion and Correction», Jieun Shin et al., *New Media & Society* 19, n.º 8, 2017.

1. *Measure What Matters: How Google, Bono, and the Gates Foundation Rock the World with OKRS*, John Doerr, 2017, p. 161.

4. «Spam Wars», Evan I. Schwartz, *MIT Technology Review*, 1 de julio de 2003.

5. «The Netflix Recommender System: Algorithms, Business Value, and Innovation», Carlos A. Gomez-Uribe y Neil Hunt, *ACM Transactions on Management Information Systems* 6, n.º 4, enero de 2016.

6. «The Amazing Ways Spotify Uses Big Data, AI and Machine Learning to Drive Business Success», Benard Marr, *Forbes*, 30 de octubre de 2017.

7. Esta cifra ha sido corroborada de forma independiente por, por ejemplo, «The Demographics of YouTube, in 5 Charts», Eric Blattberg, Digiday.com, 24 de abril de 2015.

8. El ejecutivo en cuestión era Shishir Mehrota. Todas las citas y declaraciones parafraseadas en este párrafo y el siguiente proceden de Doerr, p. 163.

9. «Beware Online ‘Filter Bubbles’», Eli Pariser, conferencia en la TED2011, Long Beach, California, 2 de mayo de 2011.

10. «The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections», Robert Epstein y Ronald E. Robertson, *Proceedings of the National Academy of Sciences* 112, n.º 33, 18 de agosto de 2015.

11. «How Google Could Rig the 2016 Election», Robert Epstein, *Politico*, 19 de agosto de 2015.

13. «How Facebook Shapes Your Feed», Will Oremus, Chris Alcantara, Jeremy B. Merrill y Artur Galocha, *The Washington Post*, 26 de octubre de 2021.

14. «Teen Vine Stars Enrage Followers by Telling Girls How to Be More Attractive», Aja Romano, *Daily Dot*, 29 de diciembre de 2013.

15. «Letter to Shareholders», Jeff Bezos, archivos de la Comisión de Valores y Bolsa, 1997.

16. Doerr, pp. 166-167.

17. *Zucked: Waking Up to the Facebook Catastrophe*, Roger McNamee, 2019, p. 41.

18. El memorando original, una instantánea fascinante de la transición del sector a una economía de la atención, puede encontrarse en su versión íntegra en «Microsoft's CEO Sent a 3,187-Word Memo and We Read It So You Don't Have To», Polly Mosendz, *The Atlantic Wire*, 10 de julio de 2014.

20. «Deep Neural Networks for YouTube Recommendations», Paul Covington, Jay Adams y Emre Sargin, *Proceedings of the 10th ACM Conference on Recommender Systems*, septiembre de 2016.

21. «Reverse Engineering the YouTube Algorithm (Part 2)», Matt Gielen, Tubfilter.com, febrero de 2017.

22. «YouTube Executives Ignored Warnings, Letting Toxic Videos Run Rampant», Mark Bergen, Bloomberg, 2 de abril de 2019.

23. Esto es así según los comentarios que hizo Neel Mohan, director de producto de YouTube, en la feria del sector Consumer Electronics Show en enero de 2018. Véase, por ejemplo, «YouTube's AI Is the Puppet Master over Most of What You Watch», Joan E. Solsman, *CNet*, 10 de enero de 2018.

24. «The Facebook Dilemma», *PBS Frontline*, 29 de octubre de 2018.

25. *Ibid.* Dijo esta frase Sandy Parakilas, exdirector de operaciones de Facebook.

26. *Ibid.* Dijo esta frase Antonio García Martínez, exdirector de producto.

27. «Exposure to Ideologically Diverse News and Opinion on Facebook», Eytan Bakshy, Solomon Messing y Lada A. Adamic, *Science* 348, n.º 6239, 7 de mayo de 2015.

29. Esta conversación se ha extraído de la sección de comentarios de «For the Next Hour I'll Be Here Answering Your Questions on Facebook», Mark Zuckerberg, Facebook.com, 30 de junio de 2015.

30. «Inside Facebook's AI Machine», Steven Levy, *Wired*, febrero de 2017.

31. «News Feed: Getting Your Content to the Right People», Adam Mosseri, presentación en el congreso Facebook F8, celebrado en San Francisco, 21 abril de 2016.

33. «I was an eng leader on Facebook's NewsFeed», Krishna Gade, Twitter, 11 de febrero de 2021. <https://twitter.com/krishnagade/status/1359908897998315521>

34. «Can Mark Zuckerberg Fix Facebook Before It Breaks Democracy?», Evan Osnos, *The New Yorker*, 17 de septiembre de 2018.

35. «TikTok and the Sorting Hat», Eugene Wei, [Eugenewei.com](https://eugenewei.com), 4 de agosto de 2020.

36. Dijo esta frase Jim McFadden, «How YouTube Drives People to the Internet's Darkest Corners», Jack Nicas, *The Wall Street Journal*, 7 de febrero de 2018.

37. «On Believing What We Remember», Ian Begg, Victoria Armour y Thérèse Kerr, *Canadian Journal of Behavioral Science* 17, 1985.

38. Todos los resultados del proyecto de investigación de Chaslot sobre las elecciones de 2016 se publicaron posteriormente en «How YouTube's A.I. Boosts Alternative Facts», Guillaume Chaslot, Medium.com, 31 de marzo de 2017.

39. «The Binge Breaker», Bianca Bosker, *The Atlantic*, noviembre de 2016.

41. «‘Our Minds Can Be Hijacked’: The Tech Insiders Who Fear a Smartphone Dystopia», Paul Lewis, *The Guardian*, 6 de octubre de 2017.

42. *Stand Out of Our Light: Freedom and Resistance in the Attention Economy*, James Williams, 2017, p. 29.

43. «The Lunatics Are Running the Asylum», Renée DiResta, conferencia pronunciada en el congreso GoogleIO, Mountain View, California, 8 de junio de 2016.

45. «Twitter's Algorithm Does Not Seem to Silence Conservatives», *The Economist*, 1 de agosto de 2020.

46. «Jack Dorsey on Twitter's Mistakes», Lauren Jackson, *The New York Times*, 7 de agosto de 2020.

47. «Microsoft's Chat Bot Was Fun for Awhile, Until It Turned into a Racist», Mathew Ingram, *Fortune*, 24 de marzo de 2016.

48. «How to Make a Bot That Isn't Racist», Sarah Jeong, Motherboard, 25 de marzo de 2016.

1. «The Psychology of Conspiracy Theories», Karen M. Douglas, Robbie M. Sutton y Aleksandra Cichocka, *Current Directions in Psychological Science* 26, n.º 6, diciembre de 2017.

2. Puede encontrarse una cronología detallada del ascenso del Pizzagate, con referencias a publicaciones individuales en «Anatomy of a Fake News Scandal», Amanda Robb, *Rolling Stone*, 16 de noviembre de 2017; y en «How the Bizarre Conspiracy Theory Behind ‘Pizzagate’ Was Spread», Craig Silverman, *BuzzFeed News*, 5 de diciembre de 2016.

3. «Trump Remains Unpopular; Voters Prefer Obama on SCOTUS Pick», Tom Jenson, Public Policy Polling, 9 de diciembre de 2016.

4. *Economist/YouGov Poll*, 20 de diciembre de 2016.

5. «YouTube Executives Ignored Warnings, Letting Toxic Videos Run Rampant», Mark Bergen, Bloomberg, 2 de abril de 2019.

6. Todas las citas de este párrafo se han extraído de *Facebook: The Inside Story*, Steven Levy, 2020, pp. 360-361.

7. «Facebook, in Cross Hairs After Election, Is Said to Question Its Influence», Mike Isaac, *The New York Times*, 12 de noviembre de 2016.

8. «Facebook Executives Shut Down Efforts to Make the Site Less Divisive», Jeff Horwitz y Deepa Seetharaman, *The Wall Street Journal*, 26 de mayo de 2020.

9. «Zuckerberg Once Wanted to Sanction Trump. Then FaceBook Wrote Rules That Accommodated Him», Elizabeth Dwoskin, Craig Timberg y Tony Romm, *The Washington Post*, 28 de junio de 2020.

10. «Zuckerberg: The Idea That Fake News on Facebook Influenced the Election Is ‘Crazy’», Casey Newton, *The Verge*, 10 de noviembre de 2016.

11. «Facebook Experiment Boosts US Voter Turnout», Zoe Corbyn, *Nature*, 2012.

12. «I want to share some thoughts on Facebook and the election», Mark Zuckerberg, Facebook.com, 12 de noviembre de 2016.

13. «Twitter Board Member: Twitter Helped Trump Win The Election», Charlie Warzel, *BuzzFeed News*, 30 de noviembre de 2016.

14. «Media in the Age of Algorithms», Tim O'Reilly, Oreilly.com, 16 de noviembre de 2016.

15. «YouTube's A.I. Was Divisive in the US Presidential Election», Guillaume Chaslot, Medium.com, 27 de noviembre de 2016.

16. El ataque de Welch y la respuesta de YouTube pueden leerse con detalle en «John Podesta Is Ready to Talk About Pizzagate», Andy Kroll, *Rolling Stone*, 9 de diciembre de 2018.

17. La metodología y los resultados que se detallan en estas páginas se publicaron por primera vez en «Emotion Shapes the Diffusion of Moralized Content in Social Networks», William J. Brady, Julian A. Wills, John T. Jost, Joshua A. Tucker y Jay J. Van Bavel, *Proceedings of the National Academy of Sciences* 114, n.º 28, 11 de julio de 2017.

18. «Twitter's Research Shows That Its Algorithm Favors Conservative Views», Emma Roth, *The Verge*, 22 de octubre de 2021.

19. «Out-Group Animosity Drives Engagement on Social Media», Steve Rathje, Jay J. Van Bavel y Sander van der Linden, *Proceedings of the National Academy of Sciences* 118, n.º 26, 29 de junio de 2021.

20. «Examining Algorithmic Amplification of Political Content on Twitter», Rumman Chowdhury y Luca Belli, blog corporativo de Twitter, 21 de octubre de 2021.

21. Twitter declaró tener 328 millones de usuarios activos mensuales en el primer trimestre de 2017. Facebook dio la cifra de 1.940 millones. YouTube no publica de forma regular datos comparables, pero en alguna ocasión ha afirmado tener más de 2.000 millones de usuarios activos mensuales. Fuentes: «Twitter's Surprising User Growth Bodes Well For 2017», Trefis Team, *Forbes*, 27 de abril de 2017. «Facebook Beats in Q1 with \$8.03B Revenue, Faster Growth to 1.94B Users», Josh Constone, *TechCrunch*, 3 de mayo de 2017.

22. La capitalización de mercado de Twitter era, en abril de 2017, de 10.680 millones de dólares y la de Facebook de 417.000 millones de dólares. A modo de comparación, la de Google era de 594.000 millones. Todas estas cifras proceden de [Macrotrends.net](https://www.macrotrends.net).

23. «Why I'm Breaking Up with Twitter», Alisyn Camerota, CNN, 12 de julio de 2017.

24. «Elliott Management's Paul Singer Seeks to Replace Twitter CEO Jack Dorsey, Source Says», Alex Sherman, CNBC, 28 de febrero de 2020.

25. «'Did We Create This Monster?' How Twitter Turned Toxic», Austin Carr y Harry McCracken, *Fast Company*, 4 de abril de 2018.

26. Tuit de Jack Dorsey (@jack), 8 de diciembre de 2018. <https://twitter.com/jack/status/1071575088695140353>

27. «Jack Dorsey's Planned Move to Africa Divides Square and Twitter Investors», Kate Rooney, CNBC, 2 de diciembre de 2019.

28. Se ha informado sobre el papel de Kaplan de forma exhaustiva e independiente, por ejemplo, en Dwoskin *et al.*; Horwitz y Seetharaman; «15 Months of Fresh Hell Inside Facebook», Nicholas Thompson y Fred Vogelstein, *Wired*, 16 de abril de 2018; y «Delay, Deny, and Deflect: How Facebook's Leaders Fought Through Crisis», Sheera Frenkel, Nicholas Confessore, Cecilia Kang, Matthew Rosenberg y Jack Nicas, *The New York Times*, 14 de noviembre de 2018.

29. «Read Mark Zuckerberg's Full 6,000-Word Letter on Facebook's Global Ambitions», Kurt Wagner y Kara Swisher, *ReCode*, 16 de febrero de 2017.

31. «The Making of a YouTube Radical», Kevin Roose, *The New York Times*, 8 de junio de 2019.

32. «Inside Facebook's A.I. Machine», Steven Levy, *Wired*, 23 de febrero de 2017.

34. «A Meta-Analytic Test of Intergroup Contact Theory», Thomas F. Pettigrew y Linda R. Tropp, *Journal of Personality and Social Psychology* 90, n.º 5, junio de 2006.

35. «Exposure to Opposing Views on Social Media Can Increase Political Polarization», Christopher A. Bail *et al.*, *Proceedings of the National Academy of Sciences* 115, n.º 37, 11 de septiembre de 2018.

36. Los sociólogos llaman a esto «efecto de homogeneidad del exogrupo». Véase, por ejemplo, «Out-Group Homogeneity Effects in Natural and Minimal Groups», Thomas M. Ostrom y Constantine Sedikides, *Psychological Bulletin* 112, n.º 3, 1992.

37. Puede encontrarse un relato exhaustivo de la falsa polarización en «The Great and Widening Divide: Political False Polarization and Its Consequences», trabajo final de máster de Victoria Parker, Universidad Wilfrid Laurier, 2018.

38. «On Trolls and Polls: How Social Media Extremists and Dissenters Exacerbate and Mitigate Political False Polarization», presentación de Victoria Parker, Universidad Wilfrid Laurier, 2019.

39. «Thinking Fast and Furious: Emotional Intensity and Opinion Polarization in Online Media», David Asker y Elias Dinas, *Public Opinion Quarterly* 83, n.º 3, otoño de 2019.

40. «The Spreading of Misinformation Online», Michela Del Vicario *et al.*, *Proceedings of the National Academy of Sciences* 113, n.º 3, 19 de enero de 2016.

41. «How Social Media Took Us from Tahrir Square to Donald Trump», Zeynep Tufekci, *MIT Technology Review*, 14 de agosto de 2018.

42. «Interview with Siva Vaidhyanathan», David Greene, National Public Radio, *Morning Edition*, 26 de diciembre de 2017.

43. «Screaming into the Void: How Outrage Is Hijacking Our Culture, and Our Minds», National Public Radio, *Hidden Brain*, 7 de octubre de 2019.

45. «Moral Outrage in the Digital Age», Molly J. Crockett, *Nature Human Behaviour* 1, 2017.

46. «Mark Warner to Facebook: Tell Me What You Know», Elaine Godfrey, *The Atlantic*, 28 de septiembre de 2017.

47. «The Facebook Dilemma», *PBS Frontline*, 29 de octubre de 2018.

49. «It's Not Misinformation. It's Amplified Propaganda», Renée DiResta, *The Atlantic*, octubre de 2021.

50. «Who's Influencing Election 2016?», William Powers, Medium.com, 23 de febrero de 2016.

51. Puede encontrarse un relato de la historia de Mackey, con detalles de las imputaciones federales contra él, en «Trump's Most Influential White Nationalist Troll Is a Middlebury Grad Who Lives in Manhattan», Luke O'Brien, *HuffPost*, 5 de abril de 2018; y «FBI Arrests Prolific Racist Twitter Troll 'Ricky Vaughn' For 2016 Election Interference», Luke O'Brien, *HuffPost*, 27 de enero de 2021.

52. «Debunking 5 Viral Images of the Migrant Caravan», Kevin Roose, *The New York Times*, 24 de octubre de 2018.

53. «Shifting Attention to Accuracy Can Reduce Misinformation Online», Gordon Pennycook *et al.*, *Nature* 592, 2021.

55. «Does YouTube's Algorithm Promote Populist Candidates in the French Presidential Elections?», Guillaume Chaslot *et al.*, Mediashift.org, 21 de abril de 2017.

56. «How an Ex-YouTube Insider Investigated Its Secret Algorithm», Paul Lewis y Erin McCormick, *The Guardian*, 2 de febrero de 2018.

58. Salvo que se indique lo contrario, todas las referencias que aparecen a continuación al estudio de Brady y Crockett a lo largo de este capítulo proceden de «The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online», William J. Brady, Molly J. Crockett y Jay J. Van Bavel, *Perspectives on Psychological Science* 15, n.º 4, junio de 2020.

59. «Attentional Capture Helps Explain Why Moral and Emotional Content Go Viral», William J. Brady, Ana P. Gantman y Jay J. Van Bavel, *Journal of Experimental Psychology* 149, n.º 4, 2020.

60. «Moral-Emotional Content and Patterns of Violent Expression and Hate Speech in Online User Comment», Jeffrey Javed y Blake Miller, documento de trabajo, abril de 2019. (Más adelante, Javed encontró trabajo en Facebook, en un equipo que optimiza la colocación de anuncios.)

1. Véase, por ejemplo, *Massacre by the River: Burmese Army Crimes Against Humanity in Tula Toli*, informe de Human Rights Watch, 19 de diciembre de 2017.

2. «Rohingya Recount Atrocities: ‘They Threw My Baby into a Fire’», Jeffrey Gettleman, *The New York Times*, 11 de octubre de 2017.

3. *Sexual Violence Against Rohingya Women and Girls in Burma*, informe de Human Rights Watch, 16 de noviembre de 2017.

4. «How Facebook's Rise Fueled Chaos and Confusion in Myanmar», Timothy McLaughlin, *Wired*, julio de 2018.

5. Todos los ejemplos de este párrafo y el siguiente se han extraído de *Hate Speech Narratives and Facilitators in Myanmar*, informe del Center for Advanced Defense Studies (C4ADS), febrero de 2016.

7. *Survey of Burma/Myanmar Public Opinion*, Center for Insights in Survey Research, 1 de abril de 2017.

8. McLaughlin, 2018.

9. «Across Myanmar, Denial of Ethnic Cleansing and Loathing of Rohingya», Hanna Beech, *The New York Times*, 24 de octubre de 2017.

10. «Genocide in the Modern Era: Social Media and the Proliferation of Hate Speech in Myanmar», Ashley Kinseth, *Tea Circle Oxford*, mayo de 2018.

11. Tuit de Max Read (@max_read), 15 de marzo de 2018 (eliminado).

12. [Tuit de Adam Mosseri \(@mosseri\)](#), 15 de marzo de 2018 (eliminado).

13. «When Is Government Web Censorship Justified? An Indian Horror Story»,
Max Fisher, *The Atlantic*, 22 de agosto de 2012.

14. «Panic Seizes India as a Region's Strife Radiates», Jim Yardley, *The New York Times*, 17 de agosto de 2012.

15. Véase, por ejemplo, «Beredar Hoax Penculikan Anak, Gelandangan Disiksa Nyaris Tewas», Fajar Eko Nugroho, *Liputan6*, 7 de marzo de 2017. «Justice by Numbers», Sana Jaffrey, *New Mandala*, 12 de enero de 2017. «The Muslim Cyber Army: What Is It and What Does It Want?», Damar Juniarto, Indonesiaatmelbourne.unimelb.edu.au, 2017.

16. «Social Media Sparked, Accelerated Egypt's Revolutionary Fire», Sam Gustin, *Wired*, 11 de febrero de 2011.

17. «Let's Design Social Media That Drives Real Change», Wael Ghonim, conferencia TED, 14 de enero de 2016.

18. «Former Facebook Exec Says Social Media Is Ripping Apart Society», James Vincent, *The Verge*, 11 de diciembre de 2017.

19. *Free Internet and the Costs to Media Pluralism: The Hazards of Zero-Rating the News*, Daniel O'Maley y Amba Kak, informe digital de CIMA, 8 de noviembre de 2018.

20. [Facebook: The Inside Story](#), Steven Levy, 2020, p. 435.

21. *Zero to One*, Thiel y Masters, 2014, p. 32.

22. «Building Global Community», Mark Zuckerberg, Facebook.com, 16 de febrero de 2017.

1. «La Otra Violencia: El Linchamiento de José Abraham y Rey David», Gema Santamaría, *Nexos*, 22 de octubre de 2015.

2. «Un ruso sobrevive a un intento de linchamiento en Cancún por insultar a los mexicanos», L.P.B., *El País*, 20 de mayo de 2017.

3. «In Frightened Mexico Town, a Mob Kills 2 Young Pollsters», Alberto Arce, Associated Press, 22 de octubre de 2015.

4. «When Fake News Kills: Lynchings in Mexico Are Linked to Viral Child-Kidnap Rumors», Patrick J. McDonnel y Cecilia Sanchez, *Los Angeles Times*, 21 de septiembre de 2018.

5. Pueden encontrarse más casos en *In the Vortex of Violence: Lynching, Extralegal Justice, and the State in Post-Revolutionary Mexico*, Gema Kloppe-Santamaría, 2020.

6. «Like. Share. Kill», Yemisi Adegoke, *BBC Africa Eye*, 12 de noviembre de 2018.

7. Puede encontrarse una explicación de la amenaza al estatus y su relevancia para la coalición de Trump, por ejemplo, en «Trump-ing Foreign Affairs: Status Threat and Foreign Policy Preferences on the Right», Rachel Marie Blum y Cristopher Sebastian Parker, *Perspectives on Politics* 17, n.º 3, agosto de 2019.

8. «The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online», William J. Brady, Molly J. Crockett y Jay J. Van Bavel, *Perspectives on Psychological Science* 15, n.º 4, junio de 2020.

9. «Eine rechtsradikale Einstellung besteht aus mehr als Fremdenhass», *Der Spiegel*, 12 de octubre de 2015.

10. «Fanning the Flames of Hate: Social Media and Hate Crime», Karsten Müller y Carlo Schwarz, *Journal of the European Economic Association* 19, n.º 4, agosto de 2021.

11. Véase, por ejemplo, «How the Parkland Teens Give Us a Glimpse of a Post-Irony Internet», Miles Klee, *Mel Magazine*, 28 de febrero de 2018.

12. «Brandstifterprozess Altena», Akantifahagen.blogspot.eu, 31 de mayo de 2016.

13. «Political Tolerance, Dogmatism, and Social Media Uses and Gratifications»,
Chamil Rathnayake y Jenifer Sunrise Winter, *Policy & Internet* 9, n.º 4, 2017.

14. «Why Narcissists Are at Risk for Developing Facebook Addiction: The Need to Be Admired and the Need to Belong», Silvia Casale y Giulia Fioravanti, *Addictive Behaviors* 76, enero de 2018.

15. «The Relationship Between Addictive Use of Social Media, Narcissism, and Self-Esteem: Findings from a Large National Survey», Cecilie Schou Andreassen, Ståle Pallesen y Mark D. Griffiths, *Addictive Behaviors* 64, enero de 2017.

16. «The Psychology of Online Political Hostility: A Comprehensive, Cross-National Test of the Mismatch Hypothesis», Alexander Bor y Michael Bang Peterson, *American Political Science Review*, 2021.

17. «Snapchat vs. Facebook: Differences in Problematic Use, Behavior Change Attempts, and Trait Social Reward Preferences», Dar Meshi, Ofir Turel, y Dan Henley, *Addictive Behaviors Reports* 12, diciembre de 2020.

18. «The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech», Eshwar Chandrasekharan *et al.*, *Proceedings of the ACM on Human-Computer Interaction* 1, noviembre de 2017.

19. Puede encontrarse un resumen accesible de los trabajos de Paluck en «Romeo & Juliet in Rwanda: How a Soap Opera Sought to Change a Nation», NPR, *Hidden Brain*, 13 de julio de 2020.

20. «Changing Climates of Conflict: A Social Network Experiment in 56 Schools», Elizabeth Levy Paluck, Hana Shepher y Peter M. Aronow, *Proceedings of the National Academy of Sciences* 113, n.º 3, 19 de enero de 2016.

21. «How Does Media Influence Social Norms? Experimental Evidence on the Role of Common Knowledge», Eric Arias, *Political Science Research and Methods* 7, n.º 3, julio de 2019.

23. «Mark Zuckerberg on Facebook's Hardest Year, and What Comes Next», Ezra Klein, *Vox*, 2 de abril de 2018.

24. «Open Letter to Mark Zuckerberg», Phandeeyar *et al.*, 5 de abril de 2018.

25. «Zuckerberg Was Called Out Over Myanmar Violence. Here's His Apology», Kevin Roose y Paul Mozur, *The New York Times*, 9 de abril de 2018.

26. *Report of Independent International Fact-Finding Mission on Myanmar*, Consejo de Derechos Humanos de las Naciones Unidas, 27 de agosto de 2018.

27. «Can Mark Zuckerberg Fix Facebook Before It Breaks Democracy?», Evan Osnos, *The New Yorker*, 17 de septiembre de 2018.

28. «Facebook Reports Fourth Quarter and Full Year 2018 Results», comunicado de prensa, Facebook Investor Relations, 30 de enero de 2019.

1. «As Germans Seek News, YouTube Delivers Far-Right Tirades», Max Fisher y Katrin Bennhold, *The New York Times*, 7 de septiembre de 2018.

2. «Revealed: Facebook hate speech exploded in Myanmar during Rohingya crisis», Libby Hogan y Michael Safi, *The Guardian*, 2 de abril de 2018.

3. «Searching for Video? Google Pushes YouTube Over Rivals», Sam Schechner, Kirsten Grind y John West, *The Wall Street Journal*, 14 de julio de 2020.

4. Más adelante Kaiser publicó sus resultados en «Public Spheres of Skepticism: Climate Skeptics' Online Comments in the German Networked Public Sphere», Jonas Kaiser, *International Journal of Communication* 11, 2017.

5. Véase, por ejemplo, «Feeding Hate with Video: A Former Alt-Right YouTuber Explains His Methods», Cade Metz, *The New York Times*, 15 de abril de 2021.

6. Los investigadores publicaron más adelante sus resultados en «The German Far-right on YouTube: An Analysis of User Overlap and User Comments», Adrian Rauchfleisch y Jonas Kaiser, *Journal of Broadcasting and Electronic Media* 64, n.º 3, 2020. La investigación apareció por primera vez en «YouTubes Algorithmen sorgen dafür, dass AfD-Fans unter sich bleiben», Adrian Rauchfleisch y Jonas Kaiser, *Vice Germany*, 22 de septiembre de 2017.

8. «The Alt-Right Can't Disown Charlottesville», Ashley Feinberg, *Wired*, 13 de agosto de 2017.

9. «Summer of Hate Challenged in Companion Civil Lawsuits», Bill Morlin, Southern Poverty Law Center, 19 de octubre de 2017.

10. «Our Extended Interview with Richard Spencer on White Nationalism», *Vice News*, 10 de diciembre de 2016.

11. [Feinberg, 2017](#).

12. «Analysis of 2017 Unite the Right Event, One Year Later», Megan Squire, Megansquire.com, agosto de 2018.

13. «The Dark Side of the Networked Public Sphere», conferencia de Jonas Kaiser para la Berkman Klein Luncheon Series de la Universidad Harvard, Cambridge, Massachusetts, 23 de enero de 2018.

14. «McInnes, Molyneux, and 4chan: Investigating Pathways to the Alt-Right», Cassie Miller, Southern Poverty Law Center, 19 de abril de 2018.

16. «Birth of a White Supremacist», Andrew Marantz, *The New Yorker*, 9 de octubre de 2017.

17. Todas las citas de «the Right Stuff» proceden de Miller, 2018.

18. Marantz, 2018.

19. Véase, por ejemplo, «Alternative Influence: Broadcasting the Reactionary Right on YouTube», Rebecca Lewis, *Data & Society*, septiembre de 2018.

20. «Jordan Peterson, Custodian of the Patriarchy», Nellie Bowles, *The New York Times*, 18 de mayo de 2018.

21. «The Alt-Right Is Recruiting Depressed People», Paris Martineau, *The Outline*, 26 de febrero de 2018.

22. *Angry White Men: Masculinity at the End of an Era*, Michael Kimmel, 2017, pp. 31-68.

23. Tuit de @SadMarshGhost, 23 de febrero de 2018. <https://twitter.com/SadMarshGhost/status/967029954016874497>

24. «On Jordan Peterson, the Alt Right and Engagement across Difference», Joel Finkelstein, *Heterodox Academy*, 18 de noviembre de 2019.

25. *Extremism*, J. M. Berger, 2018, pp. 62-89.

27. «Elliot Rodger, Isla Vista Shooting Suspect, Posted Misogynistic Video before Attack», Josh Glasstetter, Southern Poverty Law Center, 24 de mayo de 2014.

28. «The Alt-Right Is Killing People», Keegan Hankes y Alex Amend, Southern Poverty Law Center, 5 de febrero de 2018. Véase también «Understanding the Incel Community on YouTube», Kostantinos Papadamou *et al.*, *Proceedings of the ACM on Human-Computer Interaction*, octubre de 2021.

29. «The Evolution of the Manosphere across the Web», Manoel Horta Ribeiro *et al.*, *Proceedings of the Fifteenth International AAAI Conference on Web and Social Media*, 2021.

30. «The Misogynist Incel Movement Is Spreading», Lois Beckett, *The Guardian*, 3 de marzo de 2021.

31. «In the Years Since the Isla Vista Shooting, the Incel Subculture Continues to Inspire Gunmen», Jennifer Mascia, *The Trace*, 23 de mayo de 2019.

32. «How Do People Join Militias? A Leaked Oath Keepers Roster Has Answers», Ali Breland, *Mother Jones*, 27 de octubre de 2021.

33. «The Making of a YouTube Radical», Kevin Roose, *The New York Times*, 9 de junio de 2019.

34. Tuit de Chris Sacca (@sacca), 12 de enero de 2021. <https://twitter.com/sacca/status/1349055880348663808>

35. «How YouTube Built a Radicalization Machine for the Far-Right», Kelly Weill, *Daily Beast*, 19 de diciembre de 2018.

36. «From Memes to Infowars: How 75 Fascist Activists Were ‘Red-Pilled’», Robert Evans, Bellingcat.com, 11 de octubre de 2018.

37. «My Affair with the Intellectual Dark Web», Meghan Daum, Medium.com, 24 de agosto de 2018.

38. Kaiser y Rauchfleisch me proporcionaron, a lo largo de varias entrevistas entre 2019 y 2020, distintos documentos de trabajo que detallan sus métodos y resultados, además de buena parte de los datos subyacentes. Más adelante esos documentos serían la base de varios artículos y capítulos de libro revisados por pares, algunos de los cuales aún están pendientes de publicación. Véase, por ejemplo, «Birds of a Feather Get Recommended Together: Algorithmic Homophily in YouTube's Channel Recommendations in the United States and Germany», Jonas Kaiser y Adrian Rauchfleisch, *Social Media + Society* 6, n.º 4, octubre de 2020.

39. «How YouTube's Channel Recommendations Push Users to the Fringe», Craig Silverman, *BuzzFeed News*, 12 de abril de 2018.

40. «Unite the Right? How YouTube's Recommendation Algorithm Connects the U.S. Far-Right», Jonas Kaiser y Adrian Rauchfleisch, *Data & Society*, 11 de abril de 2018.

41. «Auditing radicalization pathways on YouTube», Manoel Horta Ribeiro *et al.*, *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, enero de 2020.

42. «Untrue-Tube: Monetizing Misery and Disinformation», Jonathan Albright, Medium.com, 25 de febrero 25de 2018.

43. «YouTube Executives Ignored Warnings, Letting Toxic Videos Run Rampant», Mark Bergen, Bloomberg, 2 de abril de 2019.

44. «An open letter to Mark Zuckerberg», Leonard Pozner y Veronique De La Rosa, *The Guardian*, 25 de julio de 2018.

45. «Facebook Touts Fight on Fake News, but Struggles to Explain Why InfoWars Isn't Banned», Oliver Darcy, CNN, 11 de julio de 2018.

46. «Zuckerberg: The Recode Interview», Kara Swisher, *Recode*, 8 de octubre de 2018.

47. Tuit de Jack Dorsey (@jack), 7 de agosto de 2018. <https://twitter.com/jack/status/1026984245925793792>

48. «Can Mark Zuckerberg Fix Facebook Before It Breaks Democracy?», Evan Osnos, *The New Yorker*, 17 de septiembre de 2018.

49. Tuit de Jack Dorsey (@jack), 1 de marzo de 2018. <https://twitter.com/jack/status/969234275420655616>

50. «Jack Dorsey's Push to Clean Up Twitter Stalls, Researchers Say», Deepa Seetharaman, *The Wall Street Journal*, 15 de marzo de 2020.

51. Landrum publicó sus resultados en «Differential Susceptibility to Misleading Flat Earth Arguments on YouTube», Asheley Landrum, Alex Olshansky y Othello Richards, *Media Psychology* 24, n.º 1, 2021.

52. Los sociólogos lo denominan «efecto de la verdad ilusoria». Véase, por ejemplo, «Knowledge Does Not Protect against Illusory Truth Effect», Lisa K. Fazio *et al.*, *Journal of Experimental Psychology* 144, n.º 5, octubre de 2015.

53. Este es el efecto del «conocimiento general» que se menciona en el capítulo 1. Véase Arias 2019 y Chwe 2013.

54. «Mockingbird» es una referencia a una supuesta operación de la CIA. «HRC» se refiere a Hillary Clinton. «Huma» hace referencia a Huma Abedin, que formó parte del equipo de Hillary Clinton. (*N. del e.*)

55. *Understanding Racist Activism: Theory, Methods, and Research*, Kathleen M. Blee, 2017, p. 70.

56. Este es uno de los estudios que establecen una vinculación entre la sensación de impotencia y la creencia en conspiraciones: «Beliefs in conspiracies», Marina Abalakina-Paap, *Political Psychology* 20, n.º 3, 1999.

57. «QAnon High Priest Was Just Trolling Away as Citigroup Tech Executive», William Turton y Josh Brustein, Bloomberg, 7 de octubre de 2020.

58. Este es uno de los estudios en los que se ha constatado que la creencia en conspiraciones a menudo es una forma de reafirmar sensaciones perdidas de autonomía y control: «Measuring Individual Differences in Generic Beliefs in Conspiracy Theories across Cultures: Conspiracy Mentality Questionnaire», Martin Bruder *et al.*, *Frontiers in Psychology* 4, 2013.

59. Se hablará con más detalle de este proceso en un capítulo posterior. Véase, entre muchos otros, «The Prophecies of Q», Adrienne LaFrance, *The Atlantic*, junio de 2020. «QAnon Booms on Facebook as Conspiracy Group Gains Mainstream Traction», Deepa Seetharaman, *The Wall Street Journal*, 13 de agosto de 2020. «Seven: ‘Where We Go One’» Kevin Roose *et al.*, *The New York Times*, 28 de mayo de 2020.

60. «Apple, Google Cashed in on Pizzagate-Offshoot Conspiracy App», Ben Collins y Brandy Zadrozny, NBC News, 16 de julio de 2018.

61. «How a Conspiracy Theory about Democrats Drinking Children's Blood Topped Amazon's Best-Sellers List», Kaitlyn Tiffany, Vox, 6 de marzo de 2019.

62. «QAnon Is Attracting Cops», Ali Breland, *Mother Jones*, 28 de septiembre de 2020.

63. «The Yoga World Is Riddled with Anti-Vaxxers and QAnon Believers», Cecile Guerin, *Wired UK*, 28 de enero de 2021.

64. Tuit de @_qpatriot1776_, 28 de junio de 2020, ahora eliminado por Twitter. Hay una copia disponible en un tuit de Travis View (@travis_view), 29 de junio de 2020, https://twitter.com/travis_view/status/1277634756927033345

66. «‘Carol’s Journey’: What Facebook Knew about How It Radicalized Users», Brandy Zadrozny, NBC News, 22 de octubre de 2021.

68. Tuit de Renée DiResta (@noUpside), 18 de febrero de 2018. <https://twitter.com/noupside/status/965340235251920896>

69. «Facebook Knew Calls for Violence Plagued ‘Groups’ Now Plans Overhaul», Jeff Horwitz, *The Wall Street Journal*, 31 de enero de 2021.

70. «Mark in the Middle», Casey Newton, *The Verge*, 23 de septiembre de 2020.

71. Publicación en Facebook de Dominic Fox, 8 de agosto de 2019.
<www.facebook.com/reynardine/posts/10156003037586991>.

72. «Facebook: New Zealand Attack Video Viewed 4,000 Times», BBC News, 19 de marzo de 2019.

73. «Twitter User Hacks 50,000 Printers to Tell People to Subscribe to PewDiePie», Catalin Cimpanu, ZD Net, 30 de noviembre de 2018.

74. Todas las publicaciones de usuarios proceden del hilo titulado «*ahem*», de Anonymous, 8chan, 15 de marzo de 2019. Antes en <https://www.8ch.net/poles/12916717.html>.

75. «'Shut the Site Down', Says the Creator of 8chan, a Megaphone for Gunmen», Kevin Roose, *The New York Times*, 4 de agosto de 2019.

76. Comisión Real de Investigación sobre el Ataque Terrorista contra la Mezquita de Christchurch el 15 de marzo de 2019, Comisión Real de Nueva Zelanda, diciembre de 2020.

78. «Christchurch Inquiry Report Released», Helen Sullivan, *The Guardian*, 7 de diciembre de 2020.

1. «Post No Evil», WNYC, *Radiolab*, 17 de agosto de 2018.

2. Los aspectos clave de relato de Jacob sobre el funcionamiento interno de la moderación de Facebook, así como las experiencias de sus moderadores, han sido corroborados de forma independiente. Véase «Behind the Walls of Silence», Till Krause y Hannes Grassegger, *Süddeutsche Zeitung Magazine*, 15 de diciembre de 2016. «The Low-Paid Workers Cleaning Up the Worst Horrors of the Internet», Gillian Tett, *Financial Times*, 16 de marzo de 2018. «The Secret Lives of Facebook Moderators in America», Casey Newton, *The Verge*, 25 de febrero de 2019.

3. Entrevista con Justin Osofsky, entonces vicepresidente de operaciones internacionales de Facebook, el 2 de octubre de 2018. «We, Obviously, Are Always Working to Make Sure that We Have the Right Controls and Relationships in Place. My Instinct in What You're Uncovering Here Is Probably Less an Issue at a Partner Level, at a Company to Company Level, and More of What You're Saying. Which Is Like, You Have Someone on the Front Lines That's Just Saying Something That's Inappropriate.»

4. «Startup Advice for Entrepreneurs from Y Combinator», Mark Coker, *VentureBeat*, 26 de marzo de 2007.

5. «The Hardest Lessons for Startups to Learn», conferencia de Paul Graham para la Y Combinator Startup School, abril de 2006.

6. *The Code: Silicon Valley and the Remaking of America*, Margaret O'Mara, 2019, pp. 201, 271-272.

7. *Zucked: Waking Up to the Facebook Catastrophe*, Roger McNamee, 2019, p. 48.

9. «The Education of a Libertarian», Peter Thiel, CatoUnbound.com, 13 de abril de 2009.

10. «Mouthbreathing Machiavellis Dream of a Silicon Reich», Corey Pein, *The Baffler*, 19 de mayo de 2014.

11. Un término muy utilizado que se menciona, por ejemplo, en «The History of Progress Is a History of Better Monopoly Businesses Replacing Incumbents», *Zero to One*, Thiel y Masters, 2014, p. 33.

12. Vista pública sobre el uso de las plataformas de redes sociales en las operaciones de influencia extranjera, Comisión Selecta sobre Inteligencia del Senado de Estados Unidos, 1 de agosto de 2018.

14. «A Blueprint for Content Governance and Enforcement», Mark Zuckerberg, Facebook, 15 de noviembre de 2018. <www.facebook.com/notes/751449002072082>.

16. «Facebook Overhauls News Feed in Favor of ‘Meaningful Social Interactions’», Julia Carrie Wong, *The Guardian*, 11 de enero de 2018.

17. «Five Points for Anger, One for a ‘Like’: How Facebook’s Formula Fostered Rage and Misinformation», Jeremy B. Merrill y Will Oremus, *The Washington Post*, 26 de octubre de 2021.

18. «Facebook Tried to Make Its Platform a Healthier Place. It Got Angrier Instead», Keach Hagey y Jeff Horwitz, *The Wall Street Journal*, 15 de septiembre de 2021.

20. «Whistleblower: Facebook Is Misleading the Public on Progress against Hate Speech, Violence, Misinformation», Scott Pelley, *60 Minutes*, 4 de octubre de 2021.

21. «The Welfare Effects of Social Media», Hunt Allcott *et al.*, *American Economic Review* 110, n.º 3, marzo de 2020.

22. «‘Turning point’: Mitch Fifield Flags Further Government Regulation of the Internet», Michael Koziol, *Sydney Morning Herald*, 8 de octubre de 2018.

23. «European Union Says Facebook Must Change Rules by End of 2018», Alexander Smith y Jason Abbruzzese, NBC News, 19 de septiembre de 2018.

24. «Facebook Morale Takes a Tumble Along with Stock Price», Deepa Seetharaman, *The Wall Street Journal*, 14 de noviembre de 2018.

25. «15 Months of Fresh Hell Inside Facebook», Nicholas Thompson y Fred Vogelstein, *Wired*, 16 de abril de 2019.

26. Los politólogos se refieren a la democracia mediada por unos guardianes de acceso institucionales como «democracia schumpeteriana», en honor al teórico Joseph Schumpeter. Puede encontrarse más información sobre las causas y consecuencias del declive de este sistema en *How Democracies Die*, Steven Levitsky y Daniel Ziblatt, 2018, pp. 97-117.

[27](#). «Mark Zuckerberg's Letter to Investors: The Hacker Way», CNN Money, 1 de febrero de 2012.

28. «What Happened to the Public Sphere? The Networked Public Sphere and Public Opinion Formation», Jonas Kaiser *et al.*, *Handbook of Cyber-Development, Cyber-Democracy, and Cyber-Defense*, 2016, pp. 433-459.

29. «The Yellow Vest Riots In France Are What Happens When Facebook Gets Involved with Local News», Ryan Broderick y Jules Darmanin, *BuzzFeed News*, 5 de diciembre de 2018.

30. «Demands of France's Yellow Vests», France Bleu, Opendemocracy.net, 29 de noviembre de 2018.

31. «Trends in Nonviolent Resistance and State Response: Is Violence towards Civilian-Based Movements on the Rise?», Erica Chenoweth, *Global Responsibility to Protect* 9, n.º 1, enero de 2017.

32. *Twitter and Tear Gas: The Power and Fragility of Networked Protest*, Zeynep Tufekci, 2017.

33. «Former Facebook Workers: We Routinely Suppressed Conservative News», Michael Nunez, *Gizmodo*, 9 de mayo de 2016.

34. «Republicans Press Facebook over Allegations of Bias against Conservative News», Andrea Peterson, *The Washington Post*, 11 de mayo de 2016.

35. Levy, 2020, p. 343.

36. «Three Days after Removing Human Editors, Facebook Is Already Trending Fake News», Abby Ohlheiser, *The Washington Post*, 29 de agosto de 2016.

37. El blog en cuestión se llama *Ending the Fed*. Fuente sobre su papel en las elecciones de 2016: «Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election», Robert M. Faris *et al.*, Berkman Klein Center for Internet & Society Research Paper, 2017.

38. «Inside the Two Years that Shook Facebook — and the World», Nicholas Thompson y Fred Vogelstein, *Wired*, 2 de febrero de 2018.

39. Véase, por ejemplo, «Is Facebook as Left-Leaning as Everyone Suspects?», John Brandon, Fox News, 26 de septiembre de 2016.

40. «Facebook's Mark Zuckerberg Met with Conservatives over the 'Trending' Bias Spat», Arjun Kharpal, CNBC, 19 de mayo de 2016.

41. «Twitter Is ‘Shadow Banning’ Prominent Republicans», Alex Thompson, *Vice News*, 25 de julio de 2018.

42. «Twitter's Not 'Shadow Banning' Republicans, but Get Ready to Hear that It Is», Laura Hazard Owen, NiemanLab.com, 27 de julio de 2018.

43. Tuit de Matt Gaetz (@RepMattGaetz), 25 de julio de 2018. <https://twitter.com/RepMattGaetz/status/1022224027673219072>

44. Tuit de Ronna McDaniel (@GOPChairwoman), 25 de julio de 2018. <https://twitter.com/gopchairwoman/status/1022289868620267522>

45. Tuit de Donald Trump, Jr. (@donaldtrumpjr), 25 de julio de 2018. <https://twitter.com/donaldjtrumpjr/status/1022198354468593665>

46. Pueden encontrarse resúmenes de los episodios de IBM y Microsoft, por ejemplo, en «IBM and Microsoft: Antitrust Then and Now», *CNet*, 2 de enero de 2002. Para más detalles e información sobre el legado de esos casos, véase O'Mara, 2017, pp. 341-346.

47. «Dominance Ended, I.B.M. Fights Back», Sandra Salmans, *The New York Times*, 9 de enero de 1982.

49. «What the Microsoft Antitrust Case Taught Us», Richard Blumenthal y Tim Wu, *The New York Times*, 18 de mayo de 2018.

50. «Peacetime CEO/Wartime CEO», Ben Horowitz, A16z.com, 15 de abril de 2011.

51. «To Create Culture, Start a Revolution», conferencia de Ben Horowitz en el congreso Startup Grind Global Conference, febrero de 2017.

52. «How Mark Zuckerberg Became a Wartime CEO», Casey Newton, *The Verge*, 20 de noviembre de 2018. El libro: *The Hard Thing about Hard Things*, Ben Horowitz, 2014.

53. «With Facebook at ‘War’, Zuckerberg Adopts More Aggressive Style», Deepa Seetharaman, *The Wall Street Journal*, 19 de noviembre de 2018.

54. «Delay, Deny and Deflect: How Facebook's Leaders Fought through Crisis», Sheera Frenkel, Nicholas Confessore, Cecilia Kang, Matthew Rosenberg y Jack Nicas, *The New York Times*, 14 de noviembre de 2018.

55. «Safe Space: Silicon Valley, Clubhouse, and the Cult of VC Victimhood», Zoe Schiffer y Megan Farokhmanesh, *The Verge*, 16 de julio de 2020.

56. Tuit de Balaji Srinivasan (@balajis), 1 de julio de 2020. <https://twitter.com/balajis/status/1278198087404515328>

58. Thompson y Vogelstein, 2019.

59. «Facebook Executives Shut Down Efforts to Make the Site Less Divisive», Jeff Horwitz y Deepa Seetharaman, *The Wall Street Journal*, 26 de mayo de 2020.

60. «Inside Mark Zuckerberg's Private Meetings with Conservative Pundits», Natasha Bertran y Daniel Lippman, *Politico*, 14 de octubre de 2019.

61. «Tucker Carlson: Facebook's Zuckerberg Dictating Which Political Opinions You're 'Allowed to Have'», Ian Schwartz, [Realclearpolitics.com](https://www.realclearpolitics.com), 2 de mayo de 2019.

62. «Facebook, Elections and Political Speech», Nick Clegg, <https://www.about.fb.com>, 24 de septiembre de 2019. Véase también «Facebook Says It Won't Remove Politicians' Posts for Breaking Its Rules», Adi Robertson, *The Verge*, 24 de septiembre de 2019.

63. Tuit de Sophie Zhang (@szhang_ds), 6 de junio de 2021. https://twitter.com/szhang_ds/status/1401392039414046720

64. Zhang ha contado su historia en varias ocasiones. La vez que lo hizo con mayor exhaustividad fue en «She Risked Everything to Expose Facebook. Now She's Telling Her Story», Karen Hao, *MIT Technology Review*, 29 de julio de 2021.

66. «The Case against Mark Zuckerberg: Insiders Say Facebook's CEO Chose Growth Over Safety», Elizabeth Dwoskin, Tory Newmyer y Shibani Mahtani, *The Washington Post*, 25 de octubre de 2021.

67. Let us breathe! Censorship and criminalization of online expression in Viet Nam, Informe de Amnistia Internacional, 30 de noviembre de 2020.

68. «Dissent Erupts at Facebook over Hands-Off Stance on Political Ads», Mike Isaac, *The New York Times*, 28 de octubre de 2019.

69. «Read the Letter Facebook Employees Sent to Mark Zuckerberg about Political Ads», recopilado por el *The New York Times*, 28 de octubre de 2019.

70. «I Worked on Political Ads at Facebook. They Profit by Manipulating Us»,
Yaël Eisenstat, *The Washington Post*, 4 de noviembre de 2019.

71. «Facebook Wrestles with the Features It Used to Define Social Networking», Mike Isaac, *The New York Times*, 25 de octubre de 2021.

73. El archivo, entre otros, puede verse en «Inside Facebook's Secret Rulebook for Global Political Speech», Max Fisher, *The New York Times*, 27 de diciembre de 2018.

74. «Quality Time, Brought to You by Big Tech», Arielle Pardes, *Wired*, 31 de diciembre de 2018.

75. Tuit de B. J. Fogg (@bjfogg), 11 de septiembre de 2019. <https://twitter.com/bjfogg/status/1171883692488183809>

76. «Addicted to Screens? That's Really a You Problem», Nellie Bowles, *The New York Times*, 6 de octubre de 2019.

78. «Where Silicon Valley Is Going to Get in Touch with Its Soul», Nellie Bowles, *The New York Times*, 4 de diciembre de 2017.

79. «Ex-Content Moderator Sues Facebook, Saying Violent Images Caused Her PTSD», Sandra E. Garcia, *The New York Times*, 25 de septiembre de 2018.

80. «Facebook Will Pay \$52 Million in Settlement with Moderators Who Developed PTSD on the Job», Casey Newton, *The Verge*, 12 de mayo de 2020.

1. «É horrível ser difamado pelo Bolsonaro», Débora Lopes, *Vice Portuguese*, 11 de mayo de 2013.

2. «With Amazon on Fire, Environmental Officials in Open Revolt against Bolsonaro», Ernesto Londoño y Leticia Casado, *The New York Times*, 28 de agosto de 2019.

3. «A Look at Offensive Comments by Brazil Candidate Bolsonaro», Stan Lehman, Associated Press, 29 de septiembre de 2018.

4. «URSAL, Illuminati, and Brazil's YouTube Subculture», Luiza Bandeira, Laboratorio de Investigación Digital Forense, 30 de agosto de 2018.

5. «Fast and False in Brazil», Luiza Bandeira, Laboratorio de Investigación Digital Forense, 19 de septiembre de 2018.

6. «Pesquisa Video Viewers: como os brasileiros estão consumindo vídeos em 2018», Maria Helena Marinho, Google Marketing Materials, septiembre de 2018.

7. Almeida y su equipo nos proporcionaron varios informes separados que documentaban su metodología y resultados, junto con los datos crudos subyacentes, en una serie de entrevistas hechas a principios de 2019. Compartí buena parte de ese material con YouTube antes de publicar nuestro artículo en el *New York Times*. Almeida *et al.* aún no han publicado sus investigaciones completas ni en un artículo académico formal, aunque han empleado una metodología parecida (y han llegado a conclusiones parecidas) en otros estudios revisados por pares. Véase «Auditing Radicalization Pathways on YouTube», Manoel Horta Ribeiro *et al.*, *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, enero de 2020. «Misinformation, Radicalization and Hate through the Lens of Users», Manoel Horta Ribeiro, Virgilio Almeida y Wagner Meira Jr., disertación, 30 de junio de 2020.

8. Véase, por ejemplo, «Snitch on a Teacher: Bolsonaro Win Sparks Push against ‘Indoctrination’», Dom Phillips, *The Guardian*, 30 de octubre de 2018.

9. «Education Is in the Crosshairs in Bolsonaro's Brazil», Michael Fox, *The Nation*, 12 de noviembre de 2018. Véase también «Brazil's Classrooms Become a Battleground in a Culture War», *The Economist*, 1 de diciembre de 2018.

10. Tuit de Jair Bolsonaro (@jairbolsonaro), 11 de noviembre de 2018. <https://twitter.com/jairbolsonaro/status/1061809199196368896>

11. Por ejemplo, «Brazil Replaces Far-Right Education Minister with Conspiracy Theorist», Dom Phillips, *The Guardian*, 9 de abril de 2019.

12. «Facebook Removes Pages of Brazil Activist Network before Elections», Brad Haynes, Reuters, 25 de julio de 2018.

13. «Vaccine Confidence and Hesitancy in Brazil», Amy Louise Brown *et al.*, *Cadernos de Saúde Pública* 21, septiembre de 2018.

14. «The Effects of Corrective Information About Disease Epidemics and Outbreaks: Evidence from Zika and Yellow Fever in Brazil», John M. Carey *et al.*, *Science Advances* 6, n.º 5, enero de 2020.

15. «Mapeando propagação de boatos no YouTube — Febre Amarela», Isabela Pimentel, Instituto Brasileiro de Investigación y Análisis de Datos, 8 de febrero de 2018.

16. Kaiser *et al.* compartieron sus resultados y metodología conmigo mientras estaban trabajando en el estudio, además de documentación acreditativa y datos subyacentes. Yo compartí partes relevantes de la investigación con YouTube antes de publicarlo. La mayor parte de la investigación apareció posteriormente en «Fighting Zika with Honey: An Analysis of YouTube's Video Recommendations on Brazilian YouTube», Kaiser, Rauchfleisch y Yasodara Córdova, *International Journal of Communication*, febrero de 2021.

17. La mayor parte de la metodología, los datos y los resultados referenciados en esta investigación pueden encontrarse en estos dos estudios publicados: «Analyzing and Characterizing Political Discussions in WhatsApp Public Groups», Josemar Alves Caetano *et al.*, documento de trabajo, 2018. «Characterizing Attention Cascades in WhatsApp Groups», Caetano *et al.*, *Proceedings of the 10th ACM Conference on Web Science*, junio de 2019.

18. «The Four Rs of Responsibility, Part 2: Raising authoritative content and reducing borderline content and harmful misinformation», YouTube Official Blog, 3 de diciembre de 2019.

19. Más adelante los investigadores publicaron algunos de sus resultados y métodos en «The Implications of Venturing Down the Rabbit Hole», Jonas Kaiser y Adrian Rauchfleisch, *Internet Policy Review*, 27 de junio de 2019.

20. Puede encontrarse un resumen de la investigación en «The Science of Sex Abuse», Rachel Aviv, *The New Yorker*, 6 de enero de 2013.

21. «Does Deviant Pornography Use Follow a Guttman-Like Progression?», Kathryn Seigfried-Spellar y Marcus Rogers, *Computers in Human Behavior* 29, n.º 5, septiembre de 2013.

22. Véase, entre otros, «An Integrative Review of Historical Technology and Countermeasure Usage Trends in Online Child Sexual Exploitation Material Offenders», Chad M. S. Steel *et al.*, *Forensic Science International* 33, junio de 2020. «Online Sexual Deviance, Pornography and Child Sexual Exploitation Material», Ethel Quayle, *Forensische Psychiatrie, Psychologie, Kriminologie* 14, 2020. «Prevention, Disruption and Deterrence of Online Child Sexual Exploitation and Abuse», Ethel Quayle, *ERA Forum* 21, 2020.

23. «YouTube Bans Comments on Videos of Young Children in Bid to Block Predators», Daisuke Wakabayashi, *The New York Times*, 28 de febrero de 2019.

24. «Senate Bill Targets YouTube's Kids Content amid Probe Report», Rebecca Kern, Bloomberg, 20 de junio de 2019.

25. Richard Blumenthal y Marsha Blackburn a Susan Wojcicki, 6 de junio de 2019. < <https://www.blumenthal.senate.gov/imo/media/doc/2019.06.03%20-%20YouTube%20-%20Child%20Abuse.pdf> >.

26. «Protecting Innocence in a Digital World», Vista de la Comisión de Asuntos Jurídicos del Senado, 9 de julio de 2019.

1. «W.H.O. Fights a Pandemic Besides Coronavirus: An ‘Infodemic’», Matt Richtel, *The New York Times*, 6 de febrero de 2020.

2. «Q&A: Solidifying Social Media Platforms' Role in Global Health», Devex Partnerships, Devex.com, 29 de noviembre de 2021.

3. «How WHO Is Engaging Big Tech to Fight Covid-19», Catherine Cheney, Devex.com, 14 de agosto de 2020.

4. «Facebook, Amazon, Google and More Met with WHO to Figure Out How to Stop Coronavirus Misinformation», Christina Farr y Salvador Rodriguez, CNBC, 14 de febrero de 2020.

5. «Coronavirus Cannot be Cured by Drinking Bleach or Snorting Cocaine, despite Social Media Rumors», Christina Capatides, CBS News, 9 de marzo de 2020.

6. «Coronavirus Conspiracy Video Spreads on Instagram among Black Celebrities», Brandy Zadrozny, NBC News, 13 de marzo de 2020.

7. Tuit de Brody Logan (@BrodyLogan), 16 de marzo de 2020. <https://twitter.com/BrodyLogan/status/1239406460188020736>

8. «Why Coronavirus Conspiracy Theories Flourish. And Why It Matters», Max Fisher, *The New York Times*, 8 de abril de 2020.

9. «Eight: ‘We Go All’», Kevin Roose, *The New York Times*, 4 de junio de 2020.

10. «YouTube Controls 16% of Pandemic Traffic Globally: Sandvine», Daniel Frankel, Next TV, 7 de mayo de 2020.

11. En abril de 2020, la organización activista Avaaz identificó cien publicaciones conspirativas vinculadas con el covid en Facebook con un total de 1,7 millones de comparticiones. «How Facebook Can Flatten the Curve of the Coronavirus Infodemic», Avaaz, 15 de abril de 2020.

12. Su publicación fue compartida 18.000 veces. «Fact-Checking a Facebook Conspiracy about Bill Gates, Dr. Fauci and Covid-19», Daniel Funke, Politifact, 14 de abril de 2020.

13. Su publicación recibió 90.000 comparticiones y 350.000 «me gusta». Tuit de @krisssnicolee, 7 de marzo de 2020. <https://twitter.com/krisssnicolee/status/1236309595544584192>

14. Su vídeo tuvo 4,3 millones de visualizaciones. «Cue the debunking: Two Bakersfield Doctors Go Viral with Dubious COVID Test Conclusions», Barbara Feder Ostrov, *Cal Matters*, 27 de abril de 2020.

15. «How Has Covid-19 Affected the Anti-Vaccination Movement? A Social Media Analysis», Commetric, 2 de junio de 2020.

16. «The Online Competition between Pro-and Anti-Vaccination Views», Neil F. Johnson *et al.*, *Nature* 582, mayo de 2020.

17. «The Case against Mark Zuckerberg: Insiders Say Facebook's CEO Chose Growth over Safety», Elizabeth Dwoskin, Tory Newmyer y Shibani Mahtani, *The Washington Post*, 25 de octubre de 2021.

19. «Facebook Employees Found a Simple Way to Tackle Misinformation. They ‘Deprioritized’ It after Meeting with Mark Zuckerberg, Documents Show», Billy Perrigo y Vera Bergengruen, *Time*, 10 de noviembre de 2021.

20. «Virality Project (US): Marketing Meets Misinformation», Renée DiResta e Isabella Garcia-Camargo, Stanford Internet Observatory, 26 de mayo de 2020.

22. «How the ‘Plandemic’ Movie and Its Falsehoods Spread Widely Online», Sheera Frenkel, Ben Decker y Davey Alba, *The New York Times*, 20 de mayo de 2020.

23. Cheney, agosto de 2020.

24. «The Relation between Media Consumption and Misinformation at the Outset of the SARS-CoV-2 Pandemic in the US», Kathleen Hall Jamieson y Dolores Albarracín, *Harvard Kennedy School Misinformation Review* 1, n.º 2, 2020.

25. «Coronavirus Doctors Battle Another Scourge: Misinformation», Adam Satariano, *The New York Times*, 17 de agosto de 2020.

26. Tuit de Rachel McKibbens (@rachelmckibbens), 15 de noviembre de 2021.
<https://twitter.com/RachelMcKibbens/status/1460268133302738947>

[27](#). «No, Covid-19 Vaccines Do Not ‘Shed’», Arijeta Lajka, Associated Press, 29 de abril de 2021.

28. Tuit de Rachel McKibbens (@rachelmckibbens), 11 de noviembre de 2021.
<https://twitter.com/RachelMcKibbens/status/1458881015917678594>

29. Denuncia penal, Estados Unidos de América contra Ivan Harrison Hunter, Caso 20-mj-758-hb, 27 de mayo de 2020.

30. «I Felt Hate More than Anything: How an Active Duty Airman Tried to Start a Civil War», Gisela Pérez de Acha, Kathryn Hurd y Ellie Lightfoot, *Frontline* y ProPublica, 13 de abril de 2021.

31. «Cyber Swarming, Memetic Warfare and Viral Insurgency», Alex Goldenberg y Joel Finkelstein, Network Contagion Research Institute, febrero de 2020.

33. Denuncia penal, Estados Unidos de América contra Timothy John Watson, Caso 3:20-mj-000127-RWT, 30 de octubre de 2020.

34. Denuncia penal, Estados Unidos de América contra Jessie Alexander Rush, Robert Jesus Blancas, Simon Sage Ybarra y Kenny Matthew Miksch, Caso CR-21-0121-JD, 23 de marzo de 2021.

35. «Facebook Removes Some Events Calling for Protests of Stay-at-Home Orders», Brandy Zadrozny, NBC News, 20 de abril de 2020.

36. Véase, por ejemplo, «Extremists Are Using Facebook to Organize for Civil War amid Coronavirus», Tech Transparency Project Report, 22 de abril de 2020.

37. «QAnon Booms on Facebook as Conspiracy Group Gains Mainstream Traction», Deepa Seetharaman, *The Wall Street Journal*, 13 de agosto de 2020.

38. Tuit de Nina Jankowicz (@wiczipedia), 27 de mayo de 2020. <https://twitter.com/wiczipedia/status/1265629272988954625>

39. «Facebook Bans One of the Anti-Vaccine Movement's Biggest Groups for Violating QAnon Rules», Aatif Sulleyman, *Newsweek*, 18 de noviembre de 2020.

40. «'PizzaGate' Conspiracy Theory Thrives Anew in the TikTok Era», Cecilia Kang y Sheera Frenkel, *The New York Times*, 27 de junio de 2020.

41. «Here are the QAnon Supporters Running for Congress in 2020», Alex Kaplan, Media Matters, 7 de enero de 2020 (actualizado el 27 de julio de 2021).

42. EUA contra Hunter, 2020.

43. Los detalles de las acciones de Carrillo y Justus se relatan de acuerdo con las denuncias penales federales. Véase, por ejemplo, «Alleged ‘Boogaloo’ Extremist Charged in Killing of Federal Officer during George Floyd Protest», Andrew Blankstein y Ben Collins, NBC News, 16 de junio de 2020.

44. De Acha, Hurd y Lightfoot, abril de 2021.

45. Tuit de Donald J. Trump (@realDonaldTrump), 29 de mayo de 2020. Tuit eliminado.

46. «Zuckerberg Says He's 'Struggling' with Trump's Latest Posts but Leaving Them Up», David Ingram, NBC News, 29 de mayo de 2020.

47. «Show Me State of Mind», Jelani Cobb, *This American Life* 671, 29 de marzo de 2019.

48. Tuit de Andy Mannix (@andrewmannix), 29 de mayo de 2020. <https://twitter.com/AndrewMannix/status/1266253783408930816>. Como si quisiera subrayar que sus antagonistas quizás tenían al hombre equivocado, después Mannix compartió un Premio Pulitzer por la información del *Minneapolis Star Tribune* sobre los abusos policiales que habían inspirado las protestas.

49. Tuit de Max Blumenthal (@MaxBlumenthal), 30 de mayo de 2020. Tuit eliminado.

50. «Facebook Promoted Extremism Leading to Federal Officer Dave Patrick Underwood's Murder: Lawsuit», Aaron Katersky, ABC News, 6 de enero de 2020.

51. Tuit de Lauren Tan (@sugarpirate_), 1 de junio de 2020. https://twitter.com/sugarpirate_/status/1266470996162146304

52. «How Facebook Is Undermining ‘Black Lives Matters’», *The Daily*, un pódcast del *New York Times*, 22 de junio de 2020.

53. Véase, por ejemplo, «Zuckerberg Once Wanted to Sanction Trump. Then Facebook Wrote Rules that Accommodated Him», Elizabeth Dwoskin, Craig Timberg y Tony Romm, *The Washington Post*, 28 de junio de 2020.

54. «Facebook's Tipping Point», Judd Legum, Popular Information, 27 de junio de 2020.

55. «Banning a Violent Network in the US», Facebook Newsroom, <https://www.about.fb.com>, 30 de junio de 2020.

56. «YouTube Bans Stefan Molyneux, David Duke, Richard Spencer, and More for Hate Speech», Julia Alexander, *The Verge*, 29 de junio de 2020.

57. «Reddit Bans The_Donald, Forum of Nearly 800,000 Trump Fans, over Abusive Posts», Bobby Allyn, NPR, 29 de junio de 2020.

58. «‘Facebook Is Hurting People at Scale’», Ryan Mac y Craig Silverman, *BuzzFeed News*, 23 de julio de 2020.

59. «Facebook Fails to Appease Organizers of Ad Boycott», Mike Isaac y Tiffany Hsu, *The New York Times*, 7 de julio de 2018. «When a Critic Met Facebook: ‘What They’re Doing Is Gaslighting’», Charlie Warzel, *The New York Times*, 9 de julio de 2020.

60. «Facebook's Civil Rights Audit — Final Report», Laura W. Murphy y el bufete de abogados Relman Colfax, <https://www.about.fb.com>, 8 dde julio de 2020.

63. «Nancy Pelosi Wonders How Top Facebook Employees Can ‘Look Themselves in the Mirror’ because They ‘Make Money Off Poison’», Avery Hartmans, *Business Insider*, 21 septiembre de 2020.

64. «New Steps to Protect the US Elections», Facebook Newsroom, <https://www.about.fb.com>, 3 de septiembre de 2020.

65. Tuit de @instagramcommes, 29 de octubre de 2020. mobile.<https://twitter.com/InstagramComms/status/1321957713476280320>

66. «With Election Day Looming, Twitter Imposes New Limits on U.S. Politicians — and Ordinary Users, Too», Elizabeth Dwoskin y Craig Timberg, *The Washington Post*, 9 de octubre de 2020.

67. «Additional Steps We're Taking Ahead of the 2020 US Election», Vijaya Gadde y Kayvon Beykpour, blog corporativo de Twitter, 9 de octubre de 2020.

68. «Facebook Amps Up Its Crackdown on QAnon», Sheera Frenkel, *The New York Times*, 6 de octubre de 2020. «Twitter, in Widening Crackdown, Removes over 70,000 QAnon Accounts», Kate Conger, *The New York Times*, 11 de enero de 2021.

69. «YouTube Tightens Rules on Conspiracy Videos, but Stops Short of Banning QAnon», Jennifer Elias, CNBC, 15 de octubre de 2020.

[70.](#) Social Media in 2020: Incitement, Counteraction report, 25 de noviembre de 2020.

71. «Marjorie Taylor Greene Indicated Support for Executing Prominent Democrats in 2018 and 2019 before Running for Congress», Em Steck y Andrew Kaczynski, CNN, 26 de enero de 2021.

72. «The Rise and Fall of the ‘Stop the Steal’ Facebook Group», Sheera Frenkel, *The New York Times*, 5 de noviembre de 2020.

73. «Richard Barnett, Arkansas Man Pictured Sitting at Nancy Pelosi's Desk, Arrested», The Associated Press, 8 de enero de 2021.

74. «Save Our Children Raises over \$1,000 for Nonprofit», *Westside Eagle Observer*, 28 de octubre de 2020.

75. «Capitol Attack Was Months in the Making on Facebook», *Tech Transparency Project Report*, 19 de enero de 2021.

77. «No Evidence Ballots Were Smuggled into Detroit Counting hub», Clara Hendrickson, *Detroit Free Press*, 5 de noviembre de 2020.

78. «Critics Call Gary Franchi's YouTube Channel, the Next News Network, a Hive of Conspiracy Theories. So How Has It Survived the Platform's Conspiracy Crackdown?», John Keilman, *Chicago Tribune*, 31 de octubre de 2020.

79. Tuit de @j_epp_, 4 de noviembre de 2020. Tuit eliminado.

80. Tuit de @breaking911, 5 de noviembre de 2020. Tuit eliminado.

81. «Internal Alarm, Public Shrugs: Facebook's Employees Dissect Its Election Role», Ryan Mac y Sheera Frenkel, *The New York Times*, 22 de octubre de 2021.

82. Tuit de Guillaume Chaslot (@gchaslot), 3 de diciembre de 2020. <https://twitter.com/gchaslot/status/1334615047197380610>

83. «Election Fraud Narrative», *Transparency.tube Report*, 17 de noviembre de 2020.

84. «YouTube Still Awash in False Voter Fraud Claims», *Tech Transparency Project Report*, 22 de diciembre, 2020.

85. «Trump's Far-Right Supporters Promise Violence at Today's DC Protests», Jordan Green, *Raw Story*, 6 de enero de 2021.

86. «On Far-Right Websites, Plans to Storm Capitol Were Made in Plain Sight», Laurel Wamsley, NPR, 7 de enero de 2021.

87. Green, 6 de enero de 2021.

89. «Richard Barnett Benton County Republican Rally», KNWA Fox 24, 6 de enero de 2021.

90. «How the Insurgent and MAGA Right Are Being Welded Together on the Streets of Washington D.C.», Robert Evans, Bellingcat.com, 5 de enero de 2021.

92. Se trata del acortamiento de *centipedes* ('ciempiés'), en referencia a un vídeo con imágenes de Trump que circuló por el subreddit The_Donald en el que sonaba la canción «Centipede», de la banda australiana Knife Party. (*N. del t.*)

93. «What Does 'Pedes' Mean?», naterich_stl, Reddit, 16 de marzo de 2019.

95. «Capitol Police Told to Hold Back on Riot Response on Jan. 6, Report Finds», Luke Broadwater, *The New York Times*, 13 de abril de 2021.

96. «West Virginia Lawmaker Records Himself Storming the U.S. Capitol: ‘We’re in!’», The Associated Press, 7 de enero de 2021.

97. Denuncia penal, Estados Unidos de América contra Derrick Evans, Caso 1:21-CR-337, 8 de enero de 2021.

98. Denuncia penal, Estados Unidos de América contra Jenny Cudd, Caso 1:21-cr-00068-TNM, 13 de octubre de 2021.

99. Denuncia penal, Estados Unidos de América contra Thomas Edward Caldwell, Donovan Ray Crowl y Jessica Marie Watkins, Caso 1:21-mj-00119, 19 de enero de 2021.

100. Denuncia penal, Estados Unidos de América contra Peter Francis Stager, Caso 1:21-mj-00057, 14 de enero de 2021.

[101](#). Denuncia penal, Estados Unidos de América contra Dominic Pezzola, Caso 1:21-mj-00047, 13 de enero de 2021.

102. «The Proud Boy Who Smashed a US Capitol Window Is a Former Marine», Tess Owen y Mack Lamoureux, *Vice News*, 15 de enero de 2021.

103. «The Radicalization of Kevin Greeson», Connor Sheets, ProPublica y *Birmingham News*, 15 de enero de 2021.

104. «Death of QAnon Follower at Capitol Leaves a Wake of Pain», Nicholas Bogel-Burroughs y Evan Hill, *The New York Times*, 30 de mayo de 2021.

105. «The Story of the Man Who Filmed Ashli Babbitt's Death», Samuel Benson, *Deseret News*, 11 de agosto de 2021.

106. «John Sullivan, Who Filmed Shooting of Ashli Babbitt in Capitol, Detained on Federal Charges», Robert Mackey, *The Intercept*, 14 de enero de 2021.

107. «Facebook Forced Its Employees to Stop Discussing Trump’s Coup Attempt», Ryan Mac, *BuzzFeed News*, 6 de enero de 2021.

108. «Alphabet Workers Union Statement on Yesterday's Insurrection», Alphabet Workers Union, 7 de enero de 2021.

109. Tuit de Chris Sacca (@sacca), 6 de enero de 2021. <https://twitter.com/sacca/status/1346921144859783169>

110. «Joe Biden», The Editorial Board, *The New York Times*, 17 de enero de 2020.

111. Tom Malinowski y Anna G. Eshoo a Mark Zuckerberg, 21 de enero de 2021.
<<https://www.malinowski.house.gov/files/>>.

112. Tom Malinowski y Anna G. Eshoo a Sundar Pichai y Susan Wojcicki, 21 de enero de 2021. < <https://www.malinowski.house.gov/sites/> > .

113. Publicación de Ron Watkins (@codemonkeyz), Telegram, 20 de enero de 2021.

115. Publicación de Pillow, 8kun, 20 de enero de 2021. archive.is/1G6er

116. Publicación de StartAgain, Greatawakening.win, 20 de enero de 2021.

117. Publicación de FL350, Greatawakening.win, 20 de enero de 2021.

118. Publicación de Bubba1776, Greatawakening.win, 20 de enero de 2021.

1. «Facebook to Stop Recommending Civic and Political Groups», BBC News, 28 de enero de 2021.

2. «Facebook Said It Would Stop Pushing Users to Join Partisan Political Groups. It Didn't», Leon Yin y Alfred Ng, *The Markup*, 19 de enero de 2021.

3. «Facebook Says ‘Technical Issues’ Were the Cause of Broken Promise to Congress», Alfred Ng y Leon Yin, *The Markup*, 12 de febrero de 2021.

4. «Facebook Reverses Postelection Algorithm Changes that Boosted News from Authoritative Sources», Kevin Roose, *The New York Times*, 16 de diciembre de 2020.

5. «An Update on the Georgia Runoff Elections», Sarah Schiff, <https://www.about.fb.com>, 15 de diciembre de 2020.

6. «In Georgia, Facebook's Changes Brought Back a Partisan News Feed», Corin Faife, *The Markup*, 5 de enero de 2021.

7. «Facebook Ends Ban on Political Advertising», Mike Isaac, *The New York Times*, 3 de marzo de 2021.

8. «Facebook's Sandberg Deflected Blame for Capitol Riot, but New Evidence Shows How Platform Played Role», Elizabeth Dwoskin, *The Washington Post*, 13 de enero de 2021.

9. «Banning President Trump Was the Right Decision, Says Instagram's Adam Mosseri», Nilay Patel, *The Verge*, 19 de enero de 2021.

10. «You and the Algorithm: It Takes Two to Tango», Nick Clegg, Medium.com, 31 de marzo de 2021.

11. Trending in the Wrong Direction: Social Media Platforms' Declining Enforcement of Voting Disinformation, informe de Common Cause, 2 de septiembre de 2021.

12. *Breaching the Mainstream: A National Survey of Far-Right Membership in State Legislatures*, Institute for Research and Education on Human Rights, mayo de 2022.

13. «Florida GOP pushes ‘intellectual diversity’ survey for colleges», Ana Ceballos, *The Tampa Bay Times*, 6 de abril de 2021.

14. «Voting Machine Missing after Michigan Clerk Stripped of Election Power», Jonathan Oosting, *Bridge Michigan*, 28 de octubre de 2021. «Several Interruptions from Tina Peters Caused Commissioners to Almost Throw Peters Out of Public Hearing», *Western Slope Now*, 25 de octubre de 2021.

15. «QAnon candidates are on the ballot in 26 states», Steve Reilly *et al.*, *Grid*, 12 de abril de 2022.

16. «Surgeon General Assails Tech Companies over Misinformation on Covid-19», Sheryl Gay Stolberg y Davey Alba, *The New York Times*, 15 de julio de 2021.

17. «‘They’re Killing People’: Biden Denounces Social Media for Virus Disinformation», Zolan Kanno-Youngs y Cecilia Kang, *The New York Times*, 16 de julio de 2021.

18. *Investigation of Competition in Digital Markets*, Subcomisión de la Cámara de Representantes sobre Legislación Antimonopolio, Comercial y Administrativa, 6 de octubre de 2020.

19. «Can Australia Force Google and Facebook to Pay for News?», Celina Ribeiro, *Wired*, 30 de agosto de 2020.

20. Digital News Report: Australia 2020, Centro de Investigación sobre Noticias y Medios de la Universidad de Canberra, 2020.

21. «Facebook's New Look in Australia: News and Hospitals Out, Aliens Still In», Damien Cave, *The New York Times*, 18 de febrero de 2021.

22. Tuit de Evelyn Douek (@evelyndouek), 17 de febrero de 2021. <https://twitter.com/evelyndouek/status/1362171044136710144>

23. Tuit de Sophie McNeill (@sophiemcneill), 17 de febrero de 2021. <https://twitter.com/Sophiemcneill/status/1362187114431975426>

24. Tuit de Anthony Albanese (@albomp), 17 de febrero de 2021. <https://twitter.com/AlboMP/status/1362177819304812544>

25. «Meta Says It May Shut Down Facebook and Instagram in Europe over Data-Sharing Dispute», Sam Shear, CNBC, 7 de febrero de 2022.

26. «We're Fine without Facebook, German and French Ministers Say», William Horobin y Zoe Schneeweiss, *Bloomberg News*, 7 de febrero de 2022.

27. «How Tech Workers Feel about China, AI and Big Tech's Tremendous Power», Emily Birnbaum e Issie Lapowsky, *Protocol*, 15 de marzo de 2021.

28. «Inside the Big Facebook Leak», Ben Smith, *The New York Times*, 24 de octubre de 2021.

29. «The education of Frances Haugen: How the Facebook Whistleblower Learned to Use Data as a Weapon from Years in Tech», Cat Zakrzewski y Reed Albergotti, *The Washington Post*, 11 de octubre de 2021.

30. «Whistleblower: Facebook Is Misleading the Public on Progress against Hate Speech, Violence, Misinformation», Scott Pelley, *60 Minutes*, 4 de octubre de 2021.

31. Declaración introductoria de Frances Haugen, Vista del Senado sobre Menores y Redes Sociales, 5 de octubre de 2021.

32. Zakrzewski y Albergotti, octubre de 2021.

Las redes del caos
Max Fisher

La lectura abre horizontes, iguala oportunidades y construye una sociedad mejor. La propiedad intelectual es clave en la creación de contenidos culturales porque sostiene el ecosistema de quienes escriben y de nuestras librerías. Al comprar este ebook estarás contribuyendo a mantener dicho ecosistema vivo y en crecimiento. En Grupo Planeta agradecemos que nos ayudes a apoyar así la autonomía creativa de autoras y autores para que puedan seguir desempeñando su labor. Diríjase a CEDRO (Centro Español de Derechos Reprográficos) si necesita reproducir algún fragmento de esta obra. Puede contactar con CEDRO a través de la web www.conlicencia.com o por teléfono en el 91 702 19 70 / 93 272 04 47

Título original: *The Chaos Machine. The Inside Story of How Social Media Rewired Our Minds and Our World*

© Max Fisher, 2022

Publicado en acuerdo con Little, Brown and Company, New York, New York, USA. Todos los derechos están reservados.

© Arnau Figueras Deulofeu, por la traducción, 2024

Diseño de la cubierta: Planeta Arte & Diseño Ilustraciones de la cubierta: © Shutterstock

© de esta edición: Edicions 62, S.A., 2024

Ediciones Península

Diagonal, 662-664

08034 Barcelona

edicionespeninsula@planeta.es

www.edicionespeninsula.com

Primera edición en libro electrónico (epub): enero de 2024

ISBN: 978-84-1100-236-3 (epub)

Conversión a libro electrónico: Realización Planeta

**¡Encuentra aquí tu próxima
lectura!**



¡Síguenos en redes sociales!

